JMIR Human Factors

Impact Factor (2023): 2.6
Volume 13 (2026) ISSN 2292-9495 Editor in Chief: Andre Kushniruk, BA, MSc, PhD, FACMI

Contents

Original Papers

Viewpoint

Corrigenda and Addenda

Development of the ERATbi App, a Clinical Decision Support System for Early Recovery After Traumatic Brain Injury in the ICU: Usability Study

Hsiao-Ching Yen[1*], PT, MS; I-Hui Wu[2,3,4*], MD, PhD; Wei-Ling Hsiao[5], RN, MS; Sheng-Ru Lai[6], MS; Chen-Hao Yang[7], MS; Hsien-Chi Liao[2], MD, MS; Yin-Yi Han[2,8], MD, PhD

[1]Division of Physical Therapy, Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei, Taiwan
[2]Department of Traumatology, National Taiwan University Hospital, No. 7, Chung Shan South Road, Taipei, Taiwan
[3]Department of Cardiovascular Surgery, National Taiwan University Hospital, Taipei, Taiwan
[4]Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan
[5]Department of Nursing, National Taiwan University Hospital, Taipei, Taiwan
[6]Department of Dietetics, National Taiwan University Hospital, Taipei, Taiwan
[7]Division of Respiratory Therapy, Department of Integrated Diagnostics & Therapeutics, National Taiwan University Hospital, Taipei, Taiwan
[8]Department of Anesthesiology, National Taiwan University Hospital, Taipei, Taiwan
[*]these authors contributed equally

Corresponding Author:
Yin-Yi Han, MD, PhD
Department of Traumatology, National Taiwan University Hospital, No. 7, Chung Shan South Road, Taipei, Taiwan

Abstract

Background: Early rehabilitation in neurocritical care is often underutilized due to fragmented workflows, interdisciplinary coordination challenges, and the absence of structured digital decision support. Traditional clinical decision support systems (CDSS) often address single domains and lack adaptability to the dynamic, multiprofessional workflows of intensive care units (ICUs).

Objective: To develop and evaluate the usability of the ERATbi App (Early Recovery After Traumatic Brain Injury App), a modular, tablet-based CDSS was designed to streamline early rehabilitation planning and strengthen interdisciplinary coordination for patients with moderate-to-severe traumatic brain injury (TBI) in intensive care settings.

Methods: The ERATbi app integrates four functional modules—delirium risk management, precision nutrition, stepwise early mobilization, and respiratory care for rib fractures—into a unified interface. A simulation-based usability study was conducted with 18 ICU clinicians. Evaluation metrics included System Usability Scale (SUS) scores, task completion rates, error rates, and task durations. Additional user feedback was collected via a 5-point Likert satisfaction survey and semi-structured qualitative interviews.

Results: The app demonstrated high usability (mean SUS score 83.6, SD 7.4), a 100% (18/18 participants) task completion rate, and a low error rate (4.2%). Average module completion time was 6.5 minutes, and user satisfaction was high (mean 4.7, SD 0.5). Users highlighted the value of the app's visual logic, real-time alerts, adaptive thresholds, and modular workflow integration for enhancing team coordination and decision consistency.

Conclusions: The ERATbi app demonstrated excellent usability, high user satisfaction, and clinical relevance in simulated ICU workflows. Its logic-driven, workflow- integrated design may support scalable, interdisciplinary implementation of early rehabilitation in neurocritical care settings.

(JMIR Hum Factors 2026;13:e79981) doi:10.2196/79981

KEYWORDS

brain injuries, traumatic; decision support systems; human engineering; intensive care units; rehabilitation, early; usability testing

Introduction

Epidemiology and Clinical Burden of Traumatic Brain Injury

Traumatic brain injury (TBI) is a major global public health concern and a leading cause of death and long-term disability worldwide [1]. In Taiwan, the annual incidence of hospitalized head injuries is approximately 126.1 per 100,000 population, equivalent to nearly 29,000 cases each year. Among these, an estimated 20.5% involve moderate-to-severe TBI, which is frequently associated with impaired consciousness, mechanical ventilation, and polytrauma [2].

Barriers to Early Rehabilitation in the Acute Phase

Although the benefits of early rehabilitation are increasingly supported by evidence [3,4], its implementation in the acute phase remains limited. Major barriers include safety concerns, fragmented workflows, and the absence of structured clinical decision-support systems [5-8]. Clinical hesitation often stems from hemodynamic instability, fluctuations in intracranial pressure (ICP), and the complexity of managing multiple invasive devices. These challenges force clinicians to balance therapeutic benefits against physiological risks, contributing to variability in decision-making and delayed initiation of rehabilitation.

Gaps in Multidisciplinary Integration and Decision Support

Early mobilization, nutritional therapy, delirium prevention, and respiratory management show positive impacts on outcomes of critically ill patients [9-15]. However, these interventions are typically delivered by separate disciplines with limited cross-team coordination, resulting in inconsistent implementation and delayed functional recovery. Most existing electronic medical record (EMR) systems do not provide process-based visualization or real-time interdisciplinary decision support. As a result, clinicians often depend on paper-based tools, fragmented documentation, and manual communication, further hindering timely and coordinated early rehabilitation.

Study Rationale and Objectives

To address these challenges, we developed the Early Recovery After Traumatic Brain Injury (ERATbi) app, a modular, web-based clinical decision support system (CDSS) designed to assist ICU clinicians in delivering structured early rehabilitation for patients with moderate-to-severe TBI. The system integrates four evidence-based modules: (1) delirium risk management, (2) precision nutrition therapy, (3) stepwise early mobilization, and (4) respiratory care for rib fractures. It incorporates standardized safety thresholds (eg, mean arterial pressure ≥65 mm Hg, ICP <20 mm Hg), automated alerts, and decision checkpoints to support safe and consistent clinical decision-making.

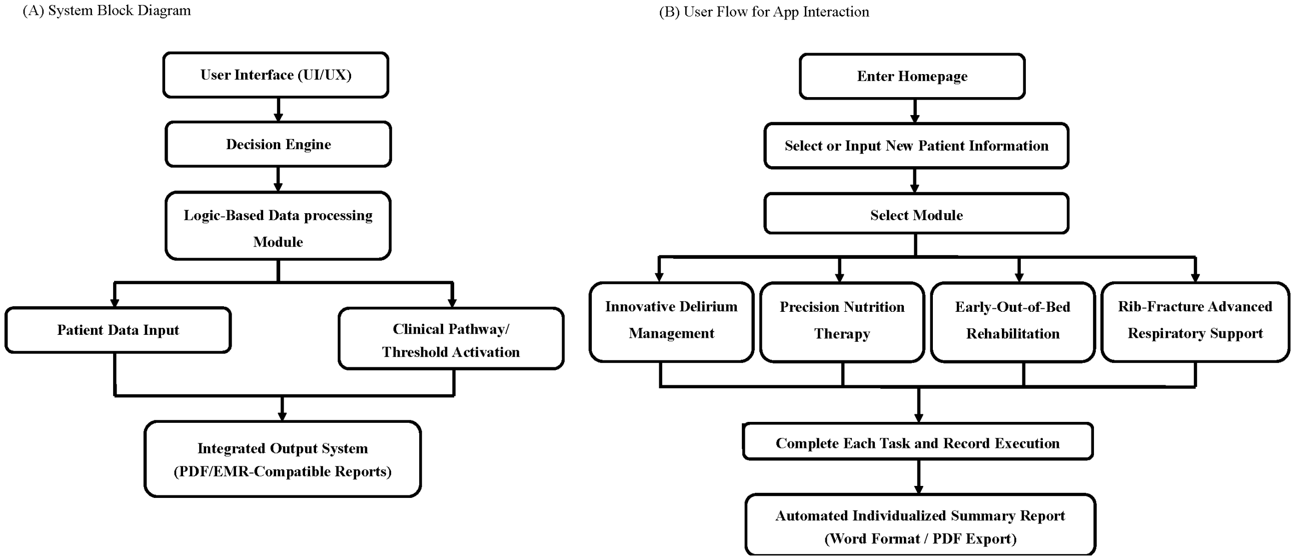
Grounded in user-centered design and interdisciplinary collaboration, this study describes the ERATbi app's theoretical framework, system architecture, and simulated clinical workflows, and evaluates its usability, clinical relevance, and potential to enhance safety, standardization, and interdisciplinary coordination in early neurocritical rehabilitation.

Methods

System Design and Development Approach

The development of the ERATbi app followed a user-centered design methodology, which is widely recommended for CDSS development in high-acuity settings due to its emphasis on workflow integration, context alignment, and stakeholder engagement [16-18] (Figure 1). A multidisciplinary co-design team—comprising intensivists, physical therapists, dietitians, critical care nurses, and software engineers—collaboratively identified workflow gaps and defined the functional requirements for a neuro-ICU–specific decision-support system. The development process consisted of three iterative stages:

Figure 1. System architecture and user workflow diagram of the ERATbi (Early Recovery After Traumatic Brain Injury) app.



Needs Assessment Phase

Direct observations and stakeholder interviews were conducted to map existing ICU rehabilitation workflows for patients with moderate-to-severe TBI and to identify barriers to timely intervention. Key gaps included inconsistent interdisciplinary coordination and the absence of real-time patient stability assessment tools.

Prototype Development Phase

Modular logic trees and interactive wireframes were constructed based on key clinical variables—such as Glasgow Coma Scale (GCS) score, time since admission, vital signs, and device constraints (eg, chest tubes, external ventricular drains). These algorithms were designed to emulate expert reasoning while supporting real-time, safety-focused decision-making.

Simulated Use Phase

Preliminary usability testing was conducted using mock clinical scenarios, enabling iterative refinement of the interface, navigation structure, and backend decision logic in response to user feedback [17].

Ethical Considerations

This study did not involve human participants or the collection of personally identifiable information. Usability evaluations were conducted exclusively in a simulated environment using standardized case scenarios developed from deidentified patient data. These data were originally obtained from an institutional review board-approved study at National Taiwan University Hospital (IRB No. 202306107RIND). The original informed consent and IRB approval covered secondary analyses of the existing data without requiring additional consent. All data used in this secondary analysis were deidentified prior to analysis, and no personally identifiable information was accessible to the study team. All participants were ICU clinicians (physicians, nurses, therapists, and dietitians) who evaluated only the app's interface and workflow. No patient contact occurred, and no clinical interventions were performed. In accordance with institutional policy, the study was deemed exempt from additional IRB review, as it did not constitute human subjects research.

Modular Architecture and Embedded Clinical Logic

The ERATbi app consists of four functional modules, each corresponding to a core domain of early neurocritical rehabilitation (Figure 2). This modular structure supports process-based clinical reasoning and enhances interdisciplinary alignment.

Figure 2. Interface storyboard of the ERATbi (Early Recovery After Traumatic Brain Injury) app demonstrating the stepwise workflow across four functional modules: precision nutrition, delirium management, stepwise early mobilization, and respiratory care.

Precision Nutrition Module

This module identifies ICU patients at risk of nutritional deficiency within the first 7 days of admission using real-time GCS and BMI calculations. Consistent with ESPEN (European Society for Clinical Nutrition and Metabolism) and ASPEN (American Society for Parenteral and Enteral Nutrition) guidelines, it stratifies patients into intervention tiers and provides individualized caloric, protein, and micronutrient recommendations [19-21]. Embedded calculators further ensure accurate risk estimation and streamline nutritional decision-making [10,22,23].

Delirium Management Module

This module incorporates the "Sweet SMART Home" protocol [24] to guide delirium risk stratification based on patient age, cognitive function, and sedation level. It emphasizes nonpharmacologic strategies such as circadian regulation and reorientation, consistent with contemporary ICU guidelines for delirium prevention [11,25].

Stepwise Early Mobilization Module

Algorithmic logic trees stratify patients according to consciousness level (GCS), physiological thresholds, and device attachments. Based on these criteria, the system generates graded activity recommendations and automated alerts for contraindicated conditions. The framework draws on established ICU mobilization models emphasizing phased recovery and functional readiness [3,26,27].

Respiratory Care for Rib Fracture Module

This module addresses comorbid thoracic trauma by supporting shared decision-making for surgical fixation and integrating electrical impedance tomography data to guide respiratory management. It includes structured prompts for pulmonary hygiene, analgesia assessment, and individualized weaning plans, aligning with ICU pain and ventilation protocols [28,29].

User Interface and Decision-Support Design

The ERATbi interface uses a process-based visual workflow that enables clinicians to navigate each module in a structured, stepwise manner. The dashboard displays real-time patient status across the four domains. Core features include auto-calculated clinical thresholds and alert mechanisms triggered by physiological or procedural contraindications.

Although full EMR integration is not yet available, the system generates structured visual summaries and consolidated recommendations to support clinical review. Data-entry fields incorporate real-time validation—such as missing-value alerts and format checks—particularly within the nutrition and mobilization modules. Guided prompts and sequential navigation are designed to reduce cognitive load and enhance data completeness during clinical use.

Simulation-Based Usability Evaluation

A simulation-based usability evaluation was conducted in the trauma ICU of a tertiary medical center. Eighteen clinicians participated, including physical therapists (n=8), ICU nurses (n=5), attending physicians (n=2), respiratory therapists (n=2), and one dietitian (n=1), all with previous experience in neurocritical care and early mobilization.

Each participant completed two standardized case vignettes reconstructed from de-identified patient records, representing moderate and severe TBI. Each vignette contained five consecutive ICU days of physiological parameters (eg, GCS, MAP [mean arterial pressure], FiO_2 [fraction of inspired oxygen], and ICP), nutritional data, and mobility progression to simulate continuous monitoring. Participants initiated a “new patient entry” and sequentially updated data to reflect daily changes, mirroring real-world decision-making workflows. Although each vignette contained data for multiple ICU days, task-level analysis was based on discrete user interactions, which were counted independently of the number of simulated days.

Case vignettes were tailored to professional scope: physical therapists emphasized activity planning, nurses focused on safety and delirium monitoring, and physicians emphasized interdisciplinary coordination. Participants used hospital-issued tablets or smartphones to complete workflows for at least two modules. All interactions were recorded via screen-capture software for subsequent analysis.

Assessment Approach

Usability was assessed using a mixed-methods approach integrating quantitative and qualitative data. Quantitative measures included the System Usability Scale (SUS) [30], task completion rate, error rate (eg, incorrect pathway selection, data-entry mistakes), and average task duration. After each simulation, participants completed a customized satisfaction questionnaire with a 5-point Likert scale (Multimedia Appendix 1) and participated in a 10 - 15-minute semi-structured interview (Multimedia Appendix 2).

Interviews were conducted face-to-face by a trained observer, audio-recorded with participant consent, and designed to explore usability facilitators, barriers, and improvement opportunities. Example prompts included the following questions. “Which parts of the interface were most intuitive or confusing?” “Were there any steps that slowed your workflow?” “What features would you modify or add to enhance clinical applicability?”

Interview recordings were transcribed verbatim and analyzed thematically following Braun and Clarke’s inductive approach. Two independent researchers coded themes related to facilitators, barriers, and design recommendations. Quantitative and qualitative findings were integrated to develop a comprehensive understanding of user experience and system performance. Sessions were monitored using validated ICU CDSS evaluation frameworks [16,17].

Methodological Rigor

To further enhance the rigor of this mixed-methods design, several strategies were implemented. For the quantitative component, task-completion metrics and SUS scores were cross-validated by two reviewers to ensure data accuracy. For the qualitative component, two researchers coded the open-ended responses using an inductive thematic approach. Coding discrepancies were discussed and resolved through consensus, and an audit trail was maintained to document analytic decisions.

Data triangulation was achieved by comparing usage metrics, task-performance patterns, and qualitative feedback. Integration of quantitative and qualitative findings followed a convergence model, whereby qualitative themes were compared with performance trends to enhance interpretive depth. Collectively, these approaches strengthened the credibility, dependability, and confirmability of the study.

Results

System Architecture Overview

Eighteen ICU professionals participated in the simulation-based usability evaluation using standardized case vignettes. All participants independently completed workflows in at least two ERATbi modules, with their interactions fully recorded for analysis. The ERATbi App integrates four core modules, Delirium Management, Precision Nutrition, Stepwise Early Mobilization, and Respiratory Care, into a unified, logic-driven clinical decision support system tailored for neurocritical rehabilitation. Each module automatically adapts recommendations based on patient-specific variables such as vital signs, level of consciousness (GCS), and device conditions. Safety features include real-time alerting, visual cue signaling deviations from physiological thresholds, and auto-generated clinical summaries. Together, these mechanisms support interdisciplinary coordination and help clinicians maintain a consistent understanding of patient status and safety priorities.

Usability Metrics

The ERATbi App demonstrated strong usability across all quantitative indicators. The System Usability Scale (SUS) yielded a mean score of 83.6 (SD 7.4), corresponding to the “excellent” usability range [31]. All participants successfully completed their assigned workflows, resulting in a 100% (18/18) task completion rate. The mean time to complete a single module was 6.5 minutes (SD 1.3). The overall error rate was 4.2% (2 errors out of 54 total interactions), primarily involving minor data-entry or selection issues.

User satisfaction, measured via a 5-point Likert scale, averaged 4.7 (SD 0.5), indicating high perceived usefulness and acceptance. Participants consistently highlighted the interface’s logical visual organization, which facilitated intuitive navigation and supported clinical reasoning, particularly beneficial for junior staff. Features such as real-time alerts, automated summaries, and visualized physiological thresholds were frequently cited as improving confidence and enhancing interdisciplinary discussions.

Qualitative Feedback

Open-ended feedback and observer field notes revealed three major themes: (1) Clarity and cognitive alignment: Participants reported that the stepwise logic structure closely aligned with their real-world ICU decision-making processes, reducing cognitive burden and supporting rapid clinical reasoning; (2) Enhanced interdisciplinary coordination: The integration of rehabilitation, nutrition, delirium, and respiratory considerations within a single interface promoted a shared mental model, enabling more coherent team discussions and planning. In this study, “shared mental model” refers to the alignment of understanding among ICU physicians, nurses, dietitians, and physical therapists regarding safety priorities, patient status, and rehabilitation goals, allowing teams to interpret clinical information consistently and coordinate care more efficiently; (3) Implementation potential: Users expressed confidence in the system’s clinical applicability and highlighted the value of future EMR integration to support handoffs, documentation consistency, and safety review workflows.

Overall, the ERATbi App demonstrated strong feasibility, high user acceptance, and operational efficiency in simulated ICU environments, supporting its potential for broader deployment in neurocritical care workflows.

Discussion

Principal Findings

This study developed and evaluated the ERATbi App, a modular CDSS designed to support early rehabilitation planning for ICU patients with moderate-to-severe TBI. In simulation testing with 18 multidisciplinary ICU clinicians, the system demonstrated excellent usability (mean SUS 83.6, SD 7.4), complete task success (100%), a low error rate (4.2%, 2/54 tasks), and high user satisfaction (mean 4.7, SD 0.5). Participants reported that the app’s structured, modular interface enhanced clarity, reduced cognitive burden, and facilitated interdisciplinary communication and real-time decision-making. Collectively, these findings indicate that ERATbi is highly feasible, operationally efficient, and well-positioned for clinical implementation in neurocritical care environments.

Overcoming the Fragmentation of Early ICU Rehabilitation Support

Prior CDSS research in ICU settings has predominantly focused on single-domain applications, including delirium screening [32,33], primary palliative care [34], rare disease diagnosis [35], fall prevention [36], or early mobility-related decision-making [37,38]. Although these systems have demonstrated benefits within their respective domains, most operate as stand-alone tools lacking interoperability, limiting their ability to support the complex, interdisciplinary workflows required in neurocritical care.

This fragmentation is compounded by EMR-integrated CDSS platforms that largely emphasize documentation or retrospective data retrieval rather than offering adaptive, real-time clinical guidance. These systems rarely adjust recommendations in response to rapidly changing physiological parameters or evolving clinical contexts, reducing their effectiveness in high-acuity decision-making.

To address these persistent gaps, the ERATbi app was conceptualized as a workflow-embedded, multi-domain CDSS specifically tailored for neurocritical care. Its design reflects the multifactorial demands of early rehabilitation in moderate-to-severe TBI—including fluctuating consciousness, invasive device management, physiological instability, and the interactions of multiple disciplines.

Evidence supporting this approach is aligned with recent studies. In 2024, Dunn et al [37] emphasized the need for standardized terminology and safety thresholds in ICU mobility tools, demonstrating high expert agreement (Content Validity Index=0.93). Similarly, in 2025, Wilson-Jene et al [38] demonstrated that algorithm-based guidance improved adherence to safety protocols and reduced mobilization delays. These studies reinforce the necessity of structured, logic-driven, and adaptive CDSS frameworks such as ERATbi.

Modular Integration and Workflow-Embedded Design in the ERATbi App

The ERATbi app was intentionally engineered as a multidomain CDSS to meet the interdisciplinary and dynamically evolving needs of neurocritical care. Unlike traditional single-domain tools, ERATbi integrates four pillars of early rehabilitation—delirium management, precision nutrition, stepwise mobilization, and respiratory care—within a unified modular interface. This architecture enables clinicians to navigate shifting consciousness levels, invasive device constraints, physiological instability, and interdisciplinary workflows with improved clarity and consistency.

Aligned with evidence-based guidelines and developed using a user-centered design methodology, ERATbi reflects contemporary recommendations that prioritize usability, workflow integration, and team-wide accessibility [39,40]. At the system's core is a decision engine that synthesizes validated physiological thresholds and individualized risk factors to generate real-time, context-aware recommendations—marking a shift from retrospective documentation tools to proactive, adaptive CDSS design.

Visual logic pathways, real-time alerts, and dynamic thresholds were incorporated to reduce cognitive load while preserving decision accuracy. Prior studies demonstrate that such features improve protocol adherence [41], reduce ICU errors [42], and enhance decision consistency during high-acuity transitions [43]. Additionally, intuitive interfaces with clearly presented thresholds have been shown to strengthen user trust, situational awareness, and adoption across multidisciplinary teams [44-46].

By embedding these technical and human-centered principles within a modular architecture, the ERATbi App provides a scalable and interoperable platform capable of standardizing early rehabilitation workflows and improving safety and communication in complex neurocritical care settings.

Clinical Implications

This study highlights several important clinical implications. First, the structured, logic-driven design of ERATbi supports consistent clinical decision-making and reduces variability across shifts and providers. Second, its visual dashboards and auto-generated summaries enhance interdisciplinary communication by providing a shared reference for physicians, nurses, dietitians, and physical therapists. Third, the app demonstrates utility as both an educational and an implementation tool, particularly for onboarding new personnel, reinforcing protocol adherence, and supporting quality improvement initiatives. Participants also noted that the modular framework could be adapted to other neurocritical populations, including stroke or post-neurosurgical patients, suggesting the system's broader clinical scalability.

Scalability and Data Integration Potential

The ERATbi app was designed for future interoperability with the hospital's HIS through HL7-FHIR–based data exchange. Structured data entered into the app (eg, GCS, MAP, ICP, FiO_2 , mobility level, and nutritional parameters) is planned to be stored using standardized FHIR Observation and Encounter resources. Physiological and biochemical data, such as vital signs, blood gas analyses, and laboratory results, are planned for future automated retrieval from the HIS within the preceding 24 - 48 hours to reduce manual input. Upon full integration, the system is planned to generate AI-assisted summary reports that consolidate the four rehabilitation modules into a single interface. These structured reports are intended to be documented in the EMR and made accessible during interdisciplinary rounds. This planned automated data exchange is expected to improve accuracy, reduce clinician workload, enhance scalability across ICUs, and enable predictive analytics by linking rehabilitation trajectories with clinical outcomes such as mobilization timing, extubation readiness, and ICU length of stay.

Limitations

This study has several limitations. First, the evaluation was conducted in a simulation-based environment rather than during real-time clinical deployment, which may affect generalizability. Although case scenarios reflected real moderate-to-severe TBI encounters, certain dynamic or emergent ICU conditions may not have been fully captured. Second, the ERATbi app was assessed as a stand-alone prototype without EMR integration, which may limit immediate scalability and clinical adoption.

Future Work

Beyond its clinical decision support function, the ERATbi system will be expanded into an educational and quality improvement platform. A planned B2C extension will provide simplified dashboards and interactive modules for patient- and family-centered education to support health literacy and post-ICU self-management. A real-world pilot study is planned at the neuro-ICU of National Taiwan University Hospital, followed by multicenter evaluations examining outcomes such as mobilization timing, delirium incidence, ICU length of stay, and safety events. Future iterations will incorporate a continuous, data-driven feedback mechanism linking clinician decisions with patient outcomes to refine decision thresholds. This iterative framework will enable adaptive modeling and sustained quality improvement in neurocritical rehabilitation.

Conclusion

The ERATbi app is a modular clinical decision support system designed to address the complex demands of early rehabilitation in neurocritical care. Simulation-based usability testing demonstrated high user acceptance, excellent usability, and strong potential to support workflow standardization, enhance interdisciplinary coordination, and promote patient safety. By integrating evidence-based clinical logic with user-centered design, the system helps to close critical gaps in ICU rehabilitation for patients with moderate-to-severe traumatic brain injury. With planned EMR integration and real-world implementation, the ERATbi app offers a scalable foundation for disseminating early rehabilitation strategies across broader neurocritical care populations.

Acknowledgments

The authors thank (1) the Department of Traumatology, National Taiwan University Hospital; (2) the National Taiwan University, College of Medicine; and (3) the Taiwanese Society of Phlebology for covering the article processing charge.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request and with appropriate institutional approvals, where applicable.

Authors' Contributions

Hsiao-Ching Yen and I-Hui Wu designed the study and contributed equally to this work. Hsiao-Ching Yen led the system development, study conception, and drafted the manuscript. I-Hui Wu contributed to the development of the clinical framework and critically revised the manuscript for important intellectual content.

Wei-Ling Hsiao, Sheng-Ru Lai, and Chen-Hao Yang contributed to module design, usability evaluation, and data interpretation.

Hsien-Chi Liao provided clinical oversight and interdisciplinary coordination.

Yin-Yi Han supervised the overall project and finalized the manuscript. All authors reviewed and approved the final manuscript and are accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

ERATbi (Early Recovery After Traumatic Brain Injury) app Usability Evaluation Questionnaire.

[DOCX File, 15 KB - humanfactors_v13i1e79981_app1.docx]

Multimedia Appendix 2

Qualitative feedback coding framework.

[DOCX File, 14 KB - humanfactors_v13i1e79981_app2.docx]

References

1. Masel BE, DeWitt DS. Traumatic brain injury: a disease process, not an event. J Neurotrauma 2010 Aug;27(8):1529-1540. [doi: 10.1089/neu.2010.1358] [Medline: 20504161]
2. Hsu IL, Li CY, Chu DC, Chien LC. An epidemiological analysis of head injuries in Taiwan. Int J Environ Res Public Health 2018 Nov 4;15(11):2457. [doi: 10.3390/ijerph15112457] [Medline: 30400380]
3. Yen HC, Han YY, Hsiao WL, et al. Functional mobility effects of progressive early mobilization protocol on people with moderate-to-severe traumatic brain injury: a pre-post intervention study. NeuroRehabilitation 2022;51(2):303-313. [doi: 10.3233/NRE-220023] [Medline: 35723117]
4. Kumar MA, Romero FG, Dharaneeswaran K. Early mobilization in neurocritical care patients. Curr Opin Crit Care 2020 Apr;26(2):147-154. [doi: 10.1097/MCC.0000000000000709] [Medline: 32068582]
5. Dirkes SM, Kozlowski C. Early mobility in the intensive care unit: evidence, barriers, and future directions. Crit Care Nurse 2019 Jun;39(3):33-42. [doi: 10.4037/ccn2019654] [Medline: 31154329]
6. Dafoe S, Chapman MJ, Edwards S, Stiller K. Overcoming barriers to the mobilisation of patients in an intensive care unit. Anaesth Intensive Care 2015 Nov;43(6):719-727. [doi: 10.1177/0310057X1504300609] [Medline: 26603796]
7. Yao Y, Dunn Lopez K, Bjarnadottir RI, et al. Examining care planning efficiency and clinical decision support adoption in a system tailoring to nurses' graph literacy: national, web-based randomized controlled trial. J Med Internet Res 2023 Aug 11;25:e45043. [doi: 10.2196/45043] [Medline: 37566456]
8. Krupp A, Steege L, Lee J, Lopez KD, King B. Supporting decision-making about patient mobility in the intensive care unit nurse work environment: work domain analysis. JMIR Nurs 2022 Sep 27;5(1):e41051. [doi: 10.2196/41051] [Medline: 36166282]
9. Adler J, Malone D. Early mobilization in the intensive care unit: a systematic review. Cardiopulm Phys Ther J 2012 Mar;23(1):5-13. [Medline: 22807649]
10. Singer P, Blaser AR, Berger MM, et al. ESPEN guideline on clinical nutrition in the intensive care unit. Clin Nutr 2019 Feb;38(1):48-79. [doi: 10.1016/j.clnu.2018.08.037] [Medline: 30348463]
11. Devlin JW, Skrobik Y, Gélinas C, et al. Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU. Crit Care Med 2018 Sep;46(9):e825-e873. [doi: 10.1097/CCM.0000000000003299] [Medline: 30113379]
12. Lewis K, Balas MC, Stollings JL, et al. A focused update to the Clinical Practice Guidelines for the Prevention and Management of Pain, Anxiety, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in adult patients in the ICU. Crit Care Med 2025 Mar 1;53(3):e711-e727. [doi: 10.1097/CCM.0000000000006574] [Medline: 39982143]

13. Prins JTH, Van Lieshout EMM, Ali-Osman F, et al. Surgical stabilization versus nonoperative treatment for flail and non-flail rib fracture patterns in patients with traumatic brain injury. Eur J Trauma Emerg Surg 2022 Aug;48(4):3327-3338. [doi: 10.1007/s00068-022-01906-1] [Medline: 35192003]
14. Franssen AJPM, Daemen JHT, Luyten JA, et al. Treatment of traumatic rib fractures: an overview of current evidence and future perspectives. J Thorac Dis 2024 Aug 31;16(8):5399-5408. [doi: 10.21037/jtd-23-1832] [Medline: 39268098]
15. Kao CC, Chen KC, Chiang XH, et al. Clinical outcomes of rib fracture stabilization and conservative treatment in a high-volume Asian trauma center: a propensity score-matched retrospective study. World J Emerg Surg 2025 May 19;20(1):40. [doi: 10.1186/s13017-025-00620-8] [Medline: 40390007]
16. Downie AS, Hancock M, Abdel Shaheed C, et al. An electronic clinical decision support system for the management of low back pain in community pharmacy: development and mixed methods feasibility study. JMIR Med Inform 2020 May 11;8(5):e17203. [doi: 10.2196/17203] [Medline: 32390593]
17. Ratwani RM, Savage E, Will A, et al. A usability and safety analysis of electronic health records: a multi-center study. J Am Med Inform Assoc 2018 Sep 1;25(9):1197-1201. [doi: 10.1093/jamia/ocy088]
18. Azadi A, García-Peñalvo FJ. Optimizing clinical decision support system functionality by leveraging specific human-computer interaction elements: insights from a systematic review. JMIR Hum Factors 2025 May 6;12:e69333. [doi: 10.2196/69333] [Medline: 40327851]
19. Díaz Chavarro BC, Romero-Saldaña M, Reveiz J, Molina-Recio G, et al. Impact of nutritional screening on mortality and intensive care unit length of stay. Front Nutr 2019;25(2):132-149. [doi: 10.3389/fnut.2025.1474039] [Medline: 31068745]
20. Singer P, Blaser AR, Berger MM, et al. ESPEN practical and partially revised guideline: clinical nutrition in the intensive care unit. Clin Nutr 2023 Sep;42(9):1671-1689. [doi: 10.1016/j.clnu.2023.07.011] [Medline: 37517372]
21. Compher C, Bingham AL, McCall M, et al. Guidelines for the provision of nutrition support therapy in the adult critically ill patient: The American Society for Parenteral and Enteral Nutrition. J Parenter Enteral Nutr 2022 Jan;46(1):12-41 [FREE Full text] [doi: 10.1002/jpen.2267]
22. Lee HY, Oh BM. Nutrition management in patients with traumatic brain injury: a narrative review. Brain Neurorehabil 2022 Mar;15(1):e4. [doi: 10.12786/bn.2022.15.e4] [Medline: 36743843]
23. Li Y, Liu C, Luo X, et al. Controlling nutritional status score and prognostic nutrition index predict the outcome after severe traumatic brain injury. Nutr Neurosci 2022 Apr;25(4):690-697. [doi: 10.1080/1028415X.2020.1804097] [Medline: 32778001]
24. Chou YH, Hsiao WL, Chen CJ, et al. Association between prosthesis contour and peri-implantitis in patients compliant with supportive periodontal therapy: a retrospective cohort study. J Prosthodont Res 2025 Jan 10;69(1):68-75. [doi: 10.2186/jpr.JPR_D_23_00236] [Medline: 39198202]
25. Pun BT, Balas MC, Barnes-Daly MA, et al. Caring for critically ill patients with the ABCDEF Bundle: results of the ICU Liberation Collaborative in over 15,000 adults. Crit Care Med 2019 Jan;47(1):3-14. [doi: 10.1097/CCM.0000000000003482] [Medline: 30339549]
26. Hodgson CL, Stiller K, Needham DM, et al. Expert consensus and recommendations on safety criteria for active mobilization of mechanically ventilated critically ill adults. Crit Care 2014 Dec 4;18(6):658. [doi: 10.1186/s13054-014-0658-y] [Medline: 25475522]
27. Zhang L, Hu W, Cai Z, et al. Early mobilization of critically ill patients in the intensive care unit: a systematic review and meta-analysis. PLoS ONE 2019;14(10):e0223185. [doi: 10.1371/journal.pone.0223185]
28. Lian KH, Yang CC, Hu FC, et al. Quality of life outcomes after surgical intervention in patients with multiple rib fractures: a prospective cohort study. Surgery 2023 Apr;173(4):1066-1071. [doi: 10.1016/j.surg.2022.12.006] [Medline: 36658082]
29. Liu HY, Lin TH, Chen KC, Hsiao WL, Hu RH, Liao HC. Comparison between non-surgical and surgical management of rib fractures in major trauma patients without brain injuries. Am J Surg 2023 Sep;226(3):350-355. [doi: 10.1016/j.amjsurg.2023.05.025] [Medline: 37263888]
30. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability Evaluation In Industry: CRC Press; 1996:189-194 URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke [accessed 2026-01-28]
31. Bangor AK, Kortum P, Miller JT. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009(3):114-123 [FREE Full text]
32. Wassenaar A, Schoonhoven L, Devlin JW, et al. Delirium prediction in the intensive care unit: comparison of two delirium prediction models. Crit Care 2018 May 5;22(1):114. [doi: 10.1186/s13054-018-2037-6] [Medline: 29728150]
33. Zhang S, Ding S, Cui W, Li X, Wei J, Wu Y. Evaluating the effectiveness of a clinical decision support system (AI-Antidelirium) to improve nurses’ adherence to delirium guidelines in the intensive care unit. Intensive Crit Care Nurs 2025 Apr;87:103933. [doi: 10.1016/j.iccn.2024.103933] [Medline: 39787945]
34. Tan A, Durbin M, Chung FR, et al. Design and implementation of a clinical decision support tool for primary palliative Care for Emergency Medicine (PRIM-ER). BMC Med Inform Decis Mak 2020 Jan 28;20(1):13. [doi: 10.1186/s12911-020-1021-7] [Medline: 31992301]
35. Schaaf J, Sedlmayr M, Sedlmayr B, Prokosch HU, Storf H. Evaluation of a clinical decision support system for rare diseases: a qualitative study. BMC Med Inform Decis Mak 2021 Feb 18;21(1):65. [doi: 10.1186/s12911-021-01435-8] [Medline: 33602191]

36. Jung H, Park HA, Lee HY. Impact of a decision support system on fall-prevention nursing practices. J Patient Saf 2023 Dec 1;19(8):525-531. [doi: 10.1097/PTS.0000000000001168] [Medline: 37922246]
37. Dunn H, da Costa Ferreira Oberfrank N, Krupp A. Preimplementation of critical care early mobility clinical decision support: a content validation study. Comput Inform Nurs 2024 Jun 1;42(6):463-469. [doi: 10.1097/CIN.0000000000001121] [Medline: 38512354]
38. Wilson-Jene H, Cowan RE, Roy Z, Pearlman J. Development of the RightWheel manual wheelchair wheel rolling resistance clinical decision support system through iterative clinician interviews. J Rehabil Assist Technol Eng 2025;12(20556683251336184):20556683251336184. [doi: 10.1177/20556683251336184] [Medline: 40291810]
39. 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):17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
40. Bright TJ, Wong A, Dhurjati R, et al. Effect of clinical decision-support systems: a systematic review. Ann Intern Med 2012 Jul 3;157(1):29-43. [doi: 10.7326/0003-4819-157-1-201207030-00450] [Medline: 22751758]
41. Nguyen AS, Yang S, Thielen BV, et al. Clinical decision support intervention and time to imaging in older patients with traumatic brain injury. J Am Coll Surg 2020 Sep;231(3):361-367. [doi: 10.1016/j.jamcollsurg.2020.05.023] [Medline: 32561447]
42. Tignanelli CJ, Shah S, Vock D, et al. A pragmatic, stepped-wedge, hybrid type II trial of interoperable clinical decision support to improve venous thromboembolism prophylaxis for patients with traumatic brain injury. Implement Sci 2024 Aug 5;19(1):57. [doi: 10.1186/s13012-024-01386-4] [Medline: 39103955]
43. Umer A, Mattila J, Liedes H, et al. A decision support system for diagnostics and treatment planning in traumatic brain injury. IEEE J Biomed Health Inform 2019 May;23(3):1261-1268. [doi: 10.1109/JBHI.2018.2842717] [Medline: 29993563]
44. Dayan PS, Ballard DW, Tham E, et al. Use of traumatic brain injury prediction rules with clinical decision support. Pediatrics 2017 Apr;139(4):e20162709. [doi: 10.1542/peds.2016-2709] [Medline: 28341799]
45. Greenberg JK, Otun A, Kyaw PT, et al. Usability and acceptability of clinical decision support based on the KIIDS-TBI tool for children with mild traumatic brain injuries and intracranial injuries. Appl Clin Inform 2022 Mar;13(2):456-467. [doi: 10.1055/s-0042-1745829] [Medline: 35477149]
46. Lang E, Neuschwander A, Favé G, et al. Clinical decision support for severe trauma patients: machine learning based definition of a bundle of care for hemorrhagic shock and traumatic brain injury. J Trauma Acute Care Surg 2022 Jan 1;92(1):135-143. [doi: 10.1097/TA.0000000000003401] [Medline: 34554136]

Abbreviations

CDSS: Clinical Decision Support System
EMR: electronic medical record
ERATbi: Early Recovery After Traumatic Brain Injury
GCS: Glasgow Coma Scale
ICP: intracranial pressure
ICU: intensive care unit
SUS: system usability scale
TBI: traumatic brain injury

Edited by C Jacob; submitted 02.Jul.2025; peer-reviewed by KH Lin, M Hellaby, M Gagnon-Roy; revised version received 13.Dec.2025; accepted 08.Jan.2026; published 06.Feb.2026.

Please cite as:
Yen HC, Wu IH, Hsiao WL, Lai SR, Yang CH, Liao HC, Han YY
Development of the ERATbi App, a Clinical Decision Support System for Early Recovery After Traumatic Brain Injury in the ICU: Usability Study
JMIR Hum Factors 2026;13:e79981
URL: https://humanfactors.jmir.org/2026/1/e79981
doi:10.2196/79981

Redesign of Bedside Supply Carts to Improve Emergency Department Workflows: Mixed Methods Participatory Design

Kat Hefter[1], MSE, JD; Sammy Dreibelbis[2], BFA; Theresa Haupt[1], MSE; Amelia McIver[1], MSE; Agnes Wang[3], MSE; Julia Dwight[4], BS, BSN; Ogechi Nwodim[5], MD; Neil Ray[5], MD

[1]Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania, 210 S. 33rd St., Suite 240, Skirkanich Hall, Philadelphia, PA, United States

[2]Department of Integrated Product Design, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA, United States

[3]Department of Data Sciences, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA, United States

[4]Emergency Department, University of Pennsylvania Health System, Philadelphia, PA, United States

[5]Department of Emergency Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Kat Hefter, MSE, JD
Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania, 210 S. 33rd St., Suite 240, Skirkanich Hall, Philadelphia, PA, United States

Abstract

Background: Emergency departments are often chaotic environments where delays can significantly impact patient care. Key items are stored in supply carts in or near patient rooms to promote efficiency and enable nurses to spend more time assisting patients. However, disorganization, lack of standardization, and lack of stocking can cause significant delays and negatively impact the quality of care.

Objective: This study utilized human-centered and participatory design to improve the workflow for supply acquisition in an emergency department.

Methods: Using a mixed methods, participatory design approach following the double diamond framework, the team worked with nursing staff and physicians in an urban emergency department to understand the root causes of frustrations with the current supply carts. Qualitative findings about bedside nursing workflows were integrated with quantitative observations of inventory and supply usage to drive a rapid-cycle prototyping process to optimize supply management in the bedside cart.

Results: A lack of clinical staffing exacerbates preexisting challenges with restocking the medical supplies in the bedside carts. This problem is compounded by the misallocation of supplies, with high-frequency items underrepresented and low-frequency items overrepresented in the bedside carts. This leads to wastage of the seldom-used supplies and a lack of access to the most used supplies. The reorganization of the cart through co-design with nursing staff sped up supply acquisition by approximately 20% overall, tripled the availability of the most important supplies, and reduced the need for restocking from once per shift to once per 3 shifts, thus producing tangible improvements even within institutional limitations.

Conclusions: A participatory design process, using human factors principles in tandem with extensive input from end users, enables improvements to stocking. Implications for practice include (1) lack of easy access to appropriate supplies negatively impacts patient care and contributes to nurse burnout and frustration, (2) human factors engineering can improve access to patient care supplies through redesigning the layout of hospital supply carts to better align with workflows, and (3) co-design with frequent collaboration from stakeholders and end users ensures that solutions address the issues that matter most in a sustainable way.

(JMIR Hum Factors 2026;13:e80861) doi:10.2196/80861

KEYWORDS

cluster analysis; emergency department; ergonomic workflow; human-centered design; inventory; medical supply cart; patient care

Introduction

In a busy emergency department (ED), every second counts. Medical professionals require access to key supplies and medication in an efficient and timely manner [1,2]. These key items are often stored in supply carts in or near patient rooms in order to promote efficiency [3-6]. However, disorganization and lack of standardization can delay patient care and negatively impact outcomes [1-3,7-9]. The lack of supplies is also a frequent cause of operational failures. Studies have shown that access to supplies is a leading factor in “time wasted” in an emergency setting [3,9,10], and a lack of familiarity with a cart system has been identified as one of the most common factors adversely affecting the quality of care in a high-stakes medical setting [11].

Human factors engineering has emerged as a dominant method to address challenges in stocking [1,2,4,7,8,12-15]. For instance, one study found that reorganizing and standardizing the contents of a resuscitation cart in a children’s hospital according to the 5S Lean design method reduced the time needed to acquire supplies by 46% [7]. Similarly, the reorganization of a hospital medication cart based on usability testing and simulation reduced the number of wasteful actions taken while searching for the proper supply [12].

While groups vary in the methods they used, one thing remains true in each case: the first step requires understanding who is most affected by the challenges caused by the existing system.

The burden of a disorganized supply system often falls heavily on nursing staff, who are already struggling with staffing shortages and heavy workloads [16,17]. Nurses report frequently having to leave rooms to search for supplies, increasing job frustration, delaying patient care, and reducing the time available to spend with patients [9,18]. One study found that more than one-third of nurses’ time is spent looking for equipment, with an annual estimated cost of US $1 million in lost time per year [10].

Nurse feedback and input are therefore a crucial component of any redesign [14,15], both to optimize any new system based on nurses’ lived experience [17] and to promote approval and uptake of a new system [1,19,20]. Nurse feedback and participation also give nurses a voice in their workplace, increasing satisfaction, improving work performance, and helping to provide an environment for safe, quality patient care [14,21].

Here, we discuss the use of a participatory human factors approach to redesign the bedside cart supply in the ED of a low-resourced urban community hospital. The aim of this mixed methods co-design intervention was to understand the challenges associated with restocking and access to patient care supplies and implement a strategy to improve reliable and efficient access to supplies for nurses in an ED. Our novel approach, which combined the use of the double diamond design framework with a participatory design, allowed us to reduce the time required to access supplies by reorganizing the carts to hold more of the most crucial supplies and minimizing wasted space.

Methods

Overview

In this paper, we report on the use of a data-driven and participatory design approach to address the supply shortage in ED bedside carts. Our approach followed the double diamond design framework [22]. The double diamond design framework is a human-centered framework that breaks a design into 2 main phases, each composed of a divergent step of discovery and ideation, followed by a convergent step distilling findings into a concise and workable next step. In addition, we relied heavily on principles of co-design [23], working closely with stakeholders at every step of the process to ensure any solutions fit within the constraints of the setting and appropriately addressed their concerns.

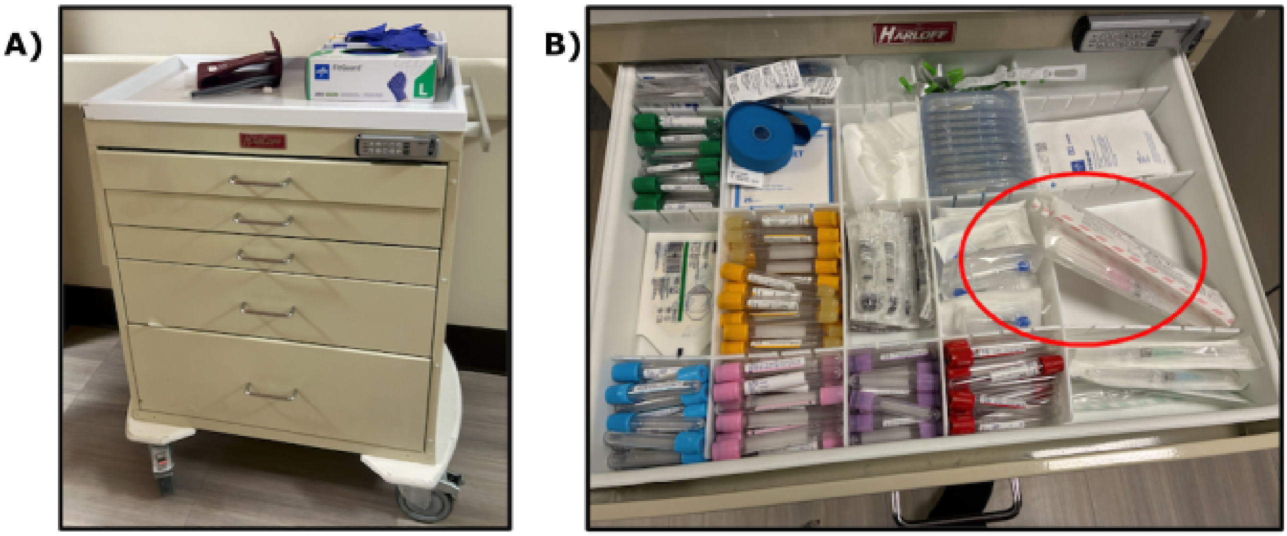
Following this framework, the design process was broken into 3 main phases. In the first phase of the double diamond, the team immersed itself in the clinical environment and engaged with stakeholders to gather information. Based on these observations, they then performed a root cause analysis to discover and define the true problems affecting supply accessibility. Finally, they underwent a co-design process with hospital staff to develop and deliver a new bedside cart prototype that successfully addressed the issue. The results are organized according to the SQUIRE (Standards for Quality Improvement Reporting Excellence) reporting guidelines for quality improvement in health care [24].

Study Background and Location

This study was conducted in the Emergency Department of the Hospital of the University of Pennsylvania’s campus at Cedar Avenue (HUP-Cedar), a low-resourced urban ED in eastern Pennsylvania. The ED serves more than 48,000 patients every year and has inpatient medical and psychiatric services, with very little specialty coverage. Children, patients who are critically injured, or patients with trauma are generally sent to other hospitals in the region, and the admission rate for patients is 10% to 13%. The department is split into a 23-bed high acuity area staffed with a 1:4 nursing to patient ratio and a 10-bed low acuity area staffed with 1 to 2 nurses.

The ED utilizes a lockable 5-drawer Harloff “bedside” cart positioned in the hallway outside each patient room. Drawers have varied heights; the top 3 are each approximately 2.5″ tall, the fourth drawer is 5.5″ tall, and the bottom drawer is 8.5″ tall. Supplies are separated by thin, $\frac{1}{8}$″ acrylic dividers in compartments that often fail to accommodate longer items such as needles and syringes (Figure 1). In addition, heavy items such as intravenous (IV) bags have deformed many of the carts over the years. Carts are often understocked given limitations in staff availability, causing nurses to search 3 or 4 carts to find necessary supplies. In addition, the underutilized supplies may linger in the cart for years well past their expiration dates. The research team was embedded within the ED from January 2025 to May 2025; the full process can be seen in Figure S1 in Multimedia Appendix 1. The team members included multiple engineers with specialties in mechanical design, product design, data science, human factors, and systems engineering and 2 emergency medicine physicians. The project also included a nurse who provided significant input and helped the embedded team interface with other members of the nursing staff.

Figure 1. Original cart system. (A) Lockable 5-drawer Harloff ‘‘bedside’’ cart positioned in the hallway of the emergency department. (B) Original layout and dividers in drawer 1 (topmost drawer) of the bedside carts. Note that needles (indicated by the red circle) are stored at a diagonal because the allocated space is too short.



Clinical Immersion and Key Stakeholder Interviews

The team began by observing users interacting with the bedside cart and interviewing key stakeholders (Multimedia Appendix 2) to understand the key challenges and concerns that they saw with the current carts, and to understand the supplies they needed the most. Interviews were semistructured [25] with a predetermined list of questions, but conversations varied according to participant interests. Interviews were conducted with multiple team members present; 1 person led the conversation, while others took detailed notes to ensure no information was missed. After 60 hours of observation, and after the insights from observation and interviewing had reached saturation [26,27], the team used a thematic analysis [28] to identify the key unmet needs (Table 1). These needs were subsequently validated through conversations with nursing staff.

Table . Sample thematic analysis of user statements.

Insight	Example statements
Unreliable or unavailable supply puts patient lives at risk.	• *[When I try to use the carts,] the things I need to save my patient’s life aren’t there.* • *I’d rather save patient lives than have a clean cart. I can clean the cart later, but I can’t bring them back.* • *I can’t leave a seizing patient to go look for oxygen masks because my cart didn’t have any.*
Stocking is time-consuming and irritating, and there is often no one to do it.	• *The techs are pulled to do other tasks like 1-on-1, and no one has time to restock the carts.* • *If we had safer ratios, I’d be more willing to stock the carts.* • *I just want my carts stocked. The rest, I can work with.* • *Anything to make things a little easier or streamlined, especially with refilling the carts, would be huge.*
Nurses do not trust the carts to have the supplies they need and often use workarounds.	• *If cart supplies are low, I’ll get a basin of the most used things and put them on a cart so I know where to find them.* • *I run from cart to cart all day, then say ‘F** it, I’m just going to the stockroom’.* • *I get things from the Clean Supply rather than the carts so I don’t waste my time looking.*
There is not a perfect consensus over what should be in the carts.	• *I’ve been going from drawer to drawer looking for socks for my patient.* • *After the emergency, that’s when we can put on socks.*
Cart contents do not match what nurses actually need.	• *A lot of things in the cart are rarely used - they’re a waste of space.* • *I really want a space for the ultrasound needles.* • *When there are things in the cart that aren’t used, they expire and we have to throw them out.*

Co-Design With Key Stakeholders and Initial Solution Selection

Focused solutions to the supply problem were addressed in a series of co-design sprints to determine how to reform the supply distribution system. A multidisciplinary group of stakeholders (including the research team, additional physicians, and nurses) underwent a series of design thinking brainstorming exercises to generate over 40 potential solutions. From there, ideas were screened for viability (whether the solution addressed the core problem), desirability (whether the stakeholders would like and use the solution), and feasibility (whether the solution would be possible given other constraints). From the session, the team identified 6 key domains—accessibility, legality, reliability, ease of restocking, capacity, and feasibility—required for a successful solution based on the unmet needs. Subsequently, screening and scoring matrices were implemented to identify the top solution [29]. Although the original solution selected was to pull out the key items used most often by nurses and put them in their own designated carts, obtaining additional carts would have become a challenge. The team sought to determine whether a similar result could be obtained merely by reallocating the proportions of various items in the existing carts to align with usage and need.

Quantitative Analysis of Current Inventory System and Item Usage

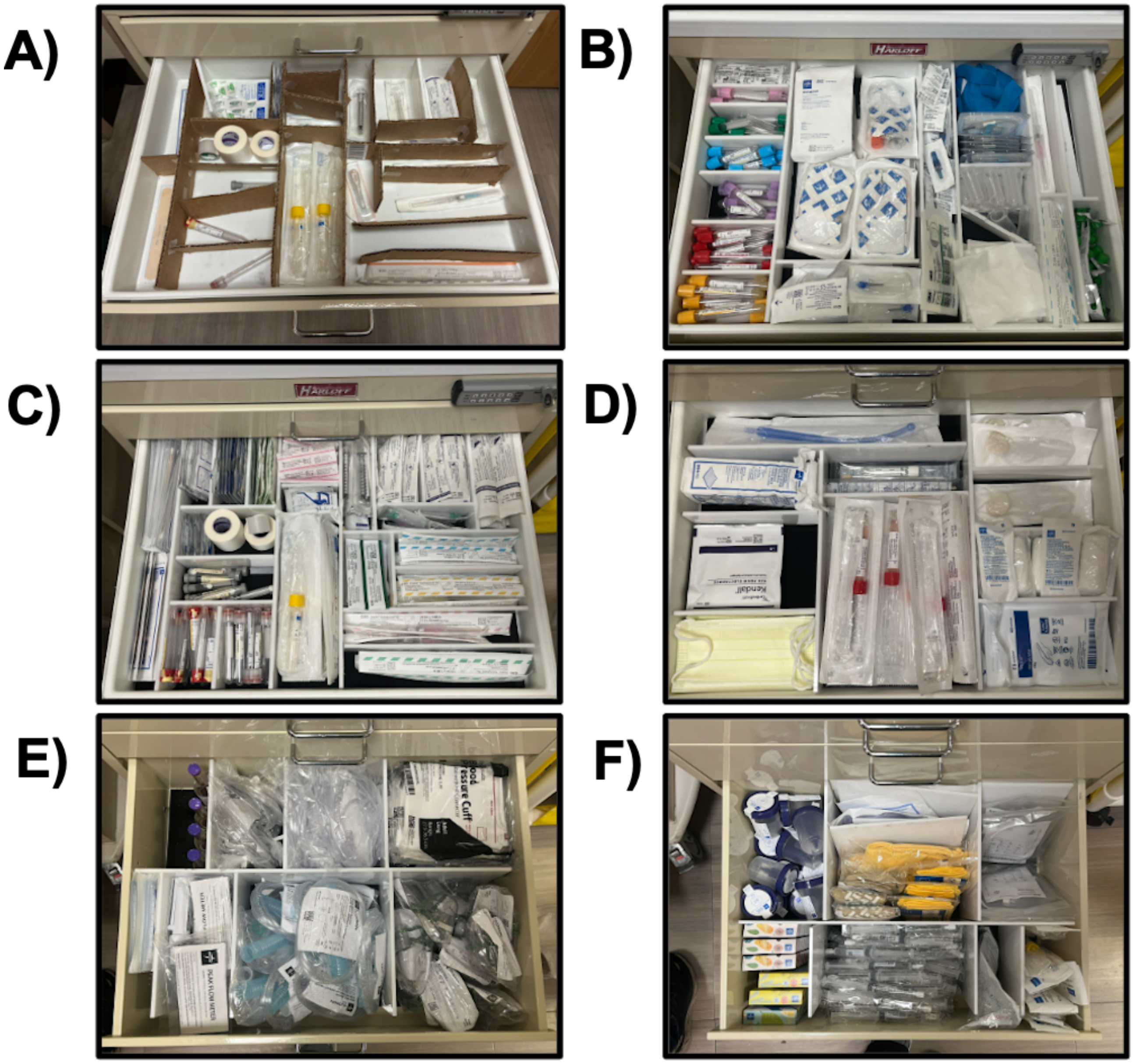
An inventory of 6 different carts throughout the ED was obtained to understand the current supply state. Maximum capacity was calculated as the highest value from each of the carts. During the observation period, the team also collected data on the supplies used to care for 17 different patients with different chief complaints (16 h of observation). This allowed the team to see supplies being used in real time and observe nursing micromovements as well as get a sense of overall inventory throughout the department.

From there, the team ran a k-means clustering analysis on the supply usage data to identify if there were any clusters of supplies that were frequently used together. While preliminary data were too limited to generate many solid clusters, the initial results using k-means with 2 clusters revealed a small cluster of high-use items and a much longer list of less frequent items. These high-use items were then compared with reports given by nurses (Multimedia Appendix 3).

Rapid Prototyping of New Drawer Organization System

Through conversation with nursing staff, the team obtained (1) a list of all of the supplies nurses wanted in the supply carts, (2) suggestions for changes in quantity for items already present, and (3) requests for layout changes to improve workflow. The team compared the wish lists to the inventory lists and eliminated unnecessary items, then integrated the qualitative suggestions with their previous quantitative results to create prototypes for drawer dividers to better allocate cart space. This was done in a “jigsaw” fashion [8]: the team built early prototypes from cardboard and tape using real supplies as a guide, then asked the nurses to move around the dividers as needed (Figure 2A). The team also asked nurses to model the movements needed to acquire items so that the new compartments had an ergonomic flow.

Figure 2. Prototype to finalized product. (A) Cardboard prototype used in an interactive design process with nurses (drawer 1); (B) finalized, fully stocked prototype for drawer 1; (C) drawer 2; (D) drawer 3; (E) drawer 4; and (F) drawer 5.



Prototypes were designed according to the following principles, all based in human factors engineering: (1) items that were used together should be located in the same drawer and in close proximity, (2) higher-frequency items should be on higher shelves to optimize body posture, (3) compartments should conform to the size and shape of the item, (4) items should fit enough supply for at least 1 full shift for a nurse, and (5) designs should minimize exterior changes to the cart to minimize relearning.

Once nurses provided feedback on the cardboard prototypes, the team incorporated those changes into new prototypes (Figure 2B-E). Team members created computer-assisted design drawings using OnShape [30] for each of the drawer bases and inserts based on the final cardboard models, then laser-cut prototypes made from ¼″ acrylic using a laser cutter.

Simulation and Evaluation of Design

To ensure that the dividers were an improvement over the original models, the team ran a variety of simulations with the help of nursing staff. The team created four scenarios that would require different combinations of supplies: (1) a patient ready for discharge, (2) a patient with fever and a recurrent headache, (3) a patient with chest and abdominal pain, and (4) a patient with multiple constitutional complaints. The number of scenarios was based on nurses' estimates of how many patients they would see in each room throughout the duration of their 12-hour shift; scenarios were intentionally left vague.

The team fully stocked both an existing cart and the prototype and asked on-shift nurses to collect the necessary supplies for each of the patient scenarios. Items removed from the cart were not replaced but instead set aside; nurses were told that if an item was not present in the cart when they needed it, they should go about their normal routine, whatever that was. This often involved searching another cart, or going to the clean supply, substituting a different item or collection of items, or foregoing that item altogether. Nurses were quasi-randomly assigned, in equal numbers, to start the simulation with either the prototype cart or the current standard to grab items from either the old or new carts first so that the time required to figure out what items to select did not bias the overall time. When all 4 scenarios were conducted with 1 cart, they would then run the same scenarios with the other cart.

Carts were not restocked between sets of simulations (“shifts”). This was to mimic real-time usage, where stockers often would not be able to resupply carts until multiple shifts had elapsed. After 3 shifts had elapsed without restocking, the carts were fully restocked, and the simulations were run again with different nurses.

Ethical Considerations

The University of Pennsylvania’s Institutional Review Board (IRB) waived the need for ethics approval and participant consent for the collection, analysis, and publication of the anonymized data for this quality improvement initiative. Informed consent was obtained verbally before participation. All data have been anonymized. No identifying information about the nurses was collected, and the team was not assessing their responses for clinical accuracy. Nurses were not compensated for their participation.

Results

Clinical Immersion and Key Stakeholder Interviews

The team spoke with a wide range of stakeholders, hoping to engage with people throughout the entire supply workflow. This included nurses, doctors, and physician assistants who used the carts; technicians who stocked the carts; and even the supply manager who oversaw stocking and ensured that the department remained in compliance with hospital policy and external regulations. The full chart of interviewees can be found in Multimedia Appendix 2.

During the observation period, the team was told that the bedside carts were fully stocked by emergency technicians and were positioned outside patient rooms so that nurses could efficiently provide patient care. In practice, however, carts were constantly understocked, and nurses frequently visited multiple carts searching for supplies. As a result of this, some nurses no longer relied on the carts and instead went to the central supply closet every time they needed something or carried a personal bucket of supplies they thought they might need, contrary to hospital policy.

After roughly 60 hours of observations and thematic analysis of statements obtained during the interviews (Table 1), it became clear that the current setup did not meet the supply and demand for current nursing staff. For example, each cart could hold only 2 IV start kits, while nurses reported using at least 1 per patient. In addition, the cart could only hold up to three 1-L bags of IV fluids, when a nurse went through an average of 5 to 10 bags per shift across all of their rooms. As a result, these key supplies were constantly in short supply and frequently depleted.

Second, while technicians were responsible for restocking these carts, they lacked sufficient time to do so amid widespread staffing shortages and a host of other, often more critical responsibilities such as sitting with patients who required continuous monitoring or helping to transport patients. Nurses could assist with restocking when time permitted, but ED shifts were often busy, and this was usually impossible. Carts often remained unstocked for several shifts, often for days on end.

In addition, while there was a photographic guide on top of the cart to assist with stocking, there was a lack of consistency in the items inside the cart. Over time, nurses had begun to stock the carts with their preferred items and formats. The clean supply room was aptly labeled with items in their respective categories—IV, wound care, comfort, respiratory, fluids, and so forth—but this system did not correspond to the organization in drawers.

Finally, the observations and interviews demonstrated the importance of buy-in for any potential solution. Nurses expressed concern that any additional work from a proposed solution would fall upon their already-overloaded shoulders and stated outright that there were certain proposals that they would simply not implement.

Co-Design With Key Stakeholders and Initial Solution Selection

The interdisciplinary team held a design sprint workshop with several nurses to generate over 40 potential solutions. Solutions ranged from developing new technologies and processes to detect when a cart was empty, to hiring a new person with the sole job of restocking carts, to altering the workflows of staff members in other departments to help assist with restocking, to even assigning each nurse their own cart to be responsible for and stock as they wished. However, solutions were limited by resources and by staffing shortages and workload. The hospital was unable to allocate money to hire a new person or replace the existing carts, and the team did not identify anyone in the environment with the extra time and ability to add restocking to their workflow. In addition, policy restrictions meant that each cart needed to be standardized. Additionally, nurses and doctors rejected any solution that involved making restocking their responsibility entirely. Given these limitations, the most feasible solution involved a reorganization of the cart system, utilizing the existing space but reconfiguring it to increase capacity for high-use items.

Quantitative Analysis of Current Inventory System and Item Usage

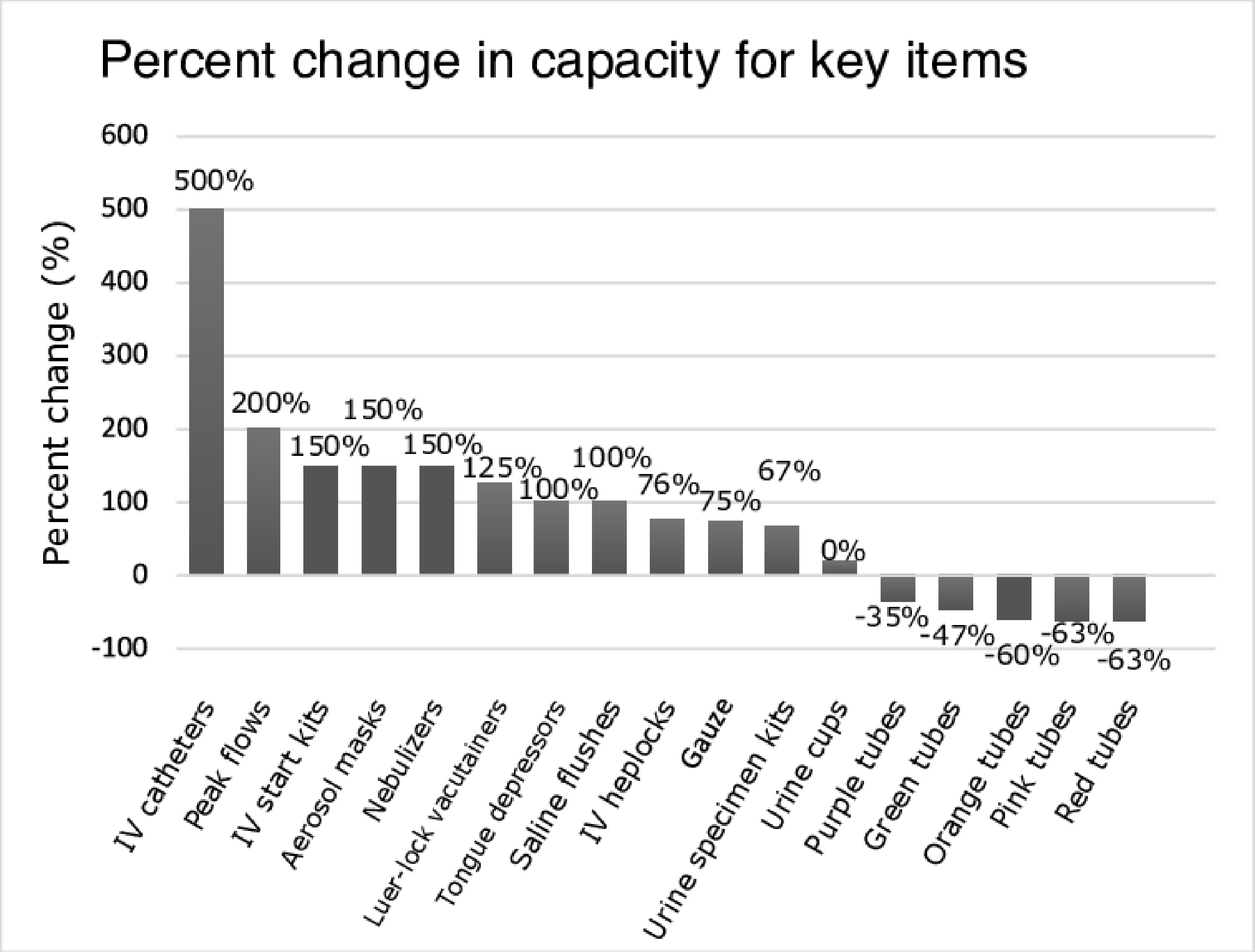
Before making any attempts to redesign the cart capacity, the team needed to understand both the original layout and typical patterns of item usage. The team thus developed an inventory to track supply usage for nurses across a total of 16 hours to capture the relative frequency of supply usage, as well as the types of supplies often used in conjunction with each other. From this information, and through consultation with nursing staff, the team was able to isolate 3 key clusters of highly used supplies: IV, diagnostics, and respiratory supplies. The contents of each of these clusters can be found in Multimedia Appendix 3. They also obtained an approximate estimate of the amount of each of these supplies needed per day. IV start kits were the most highly used item, as they were used with nearly every patient. Along with the kits, nurses also used flushes, heplocks, and vacutainers.

Rapid Prototyping of New Drawer Organization System

After undergoing the rapid prototyping process described above and shown in Figure 2, the new design provided a marked improvement over the old model. The layouts for the old versus original designs can be seen in Figure 3, and the difference in capacity between the 2 designs can be seen in Figure 4 (key supplies only; full list in Multimedia Appendix 4). Specifically, for the IV kits, the new carts comfortably held 9 kits, while previous models held only 2 or 3—a 200% increase for the most highly used supply. Co-occurring supplies, such as the flushes (100% increase), vacutainers (125% increase), and heplocks (75% increase), similarly increased in capacity. Given the total number of carts on the floor, this meant that the number of IV kits stored in the carts jumped from approximately 100 to approximately 300—enough to make it through multiple shifts without restocking. The new cart design similarly doubled the size of spaces allocated for key respiratory supplies such as peak flow meters (200% increase), nebulizers (150% increase), and aerosol masks (150% increase).

Figure 3. Drawer 1 original versus new layout. Colored boxes highlight key changes that were made. Blue: Needles are no longer stored at a diagonal with no decrease in capacity. Purple: Hep-lock capacity increased by 75%. Green: Intravenous start kit capacity increased by 200%. Orange: Storing all tubes in 1 column makes for a more ergonomic workflow; capacity decreased to avoid wastage.

Figure 4. Cart redesign corrects supply allocation. After redesign, critical items like those used with intravenous (IV) or for urine collection and testing increased in capacity, while overrepresented test tubes decreased in capacity.



In many cases, this was achieved by removing low-frequency items that were rarely used and could be more effectively stored in the central supply closet. In other cases, it was achieved by ensuring each compartment fit the size of the tool it was meant to contain, eliminating wasted space and reducing the chances of needles being damaged or bent by attempting to force them into the cart. The new design also allowed space for the addition of other supplies, such as a speculum, that were highly sought after by physicians and nurses and could not fit in the cart previously. The design met with approval from nursing and administrative staff at all stages of the process.

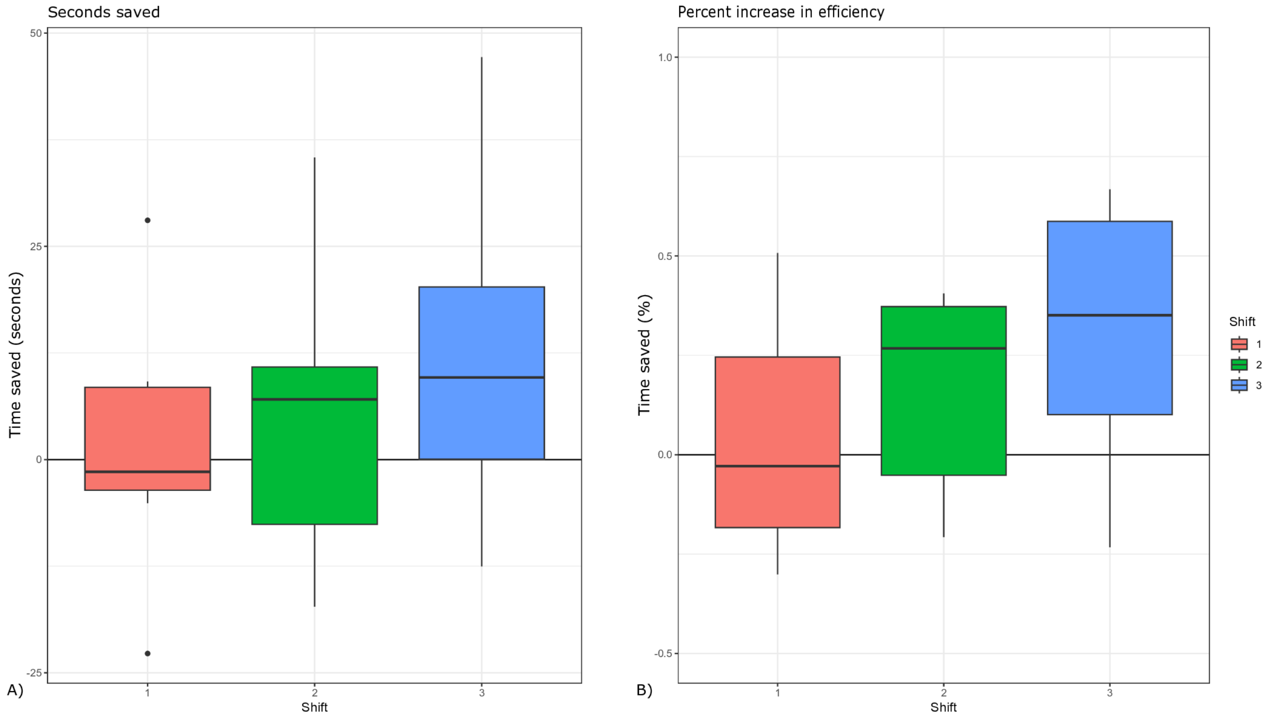
In addition, the finalized model was both sturdier and less prone to breakage than the original model. This was the result of several iterations: the first draft, which used an acrylic solvent, proved unstable and snapped upon construction, so the team reconfigured the designs to use a press-fit layout instead. With this approach, drawers were easy to assemble and sturdy enough to withstand constant use.

Simulation and Evaluation of Design

Once the prototypes were completed, they were tested for practical usage. After completing the simulation described above, the novel carts outperformed their older counterparts. On average, nurses were able to acquire needed supplies from the new cart nearly 6 seconds faster (21% faster) than they were able to get supplies from the old cart (2-tailed paired *t* test, time to acquire item from old cart=38.2 seconds; time to acquire item from new cart=32.4 seconds; degrees of freedom=23*;* *P*=.097). The lack of statistical significance was likely a limitation of low sample size; in addition, this simulation likely overestimated the time required to get samples from the newer cart because nurses were unfamiliar with the new configuration.

In addition, the time saved only increased as the carts went longer intervals without needing to be restocked. During the first shift, the new cart only saved about 1.5 seconds on average. During the second shift, the new cart saved an average of 4.7 seconds. On the third shift, the cart saved an average of 11.3 seconds (Figure 5). This time is a direct reflection of an improvement in capacity. While the old cart ran out of IV start kits by the third patient, forcing nurses to search for kits in nearby carts for every subsequent patient, the new cart contained all of the needed supplies for every patient. On average, a trip from a cart to the clean supply room (the preferred solution for nurses who did not trust that supply carts contain the proper equipment) took 1 minute; these trips were entirely eliminated by the use of the new prototype.

Figure 5. Time savings as a function of new cart. (A) The new cart saved an average of 5.8 seconds per simulated patient, and the time saved increased as the cart went longer without being restocked. (B) As the cart went longer without restocking, the percentage of time saved while accessing needed supplies increased, from about even when both were fully stocked, to 40% after 2 shifts. Boxplots show distribution of data (n=8 per shift). Sample size was too low to assess statistical significance.



Furthermore, the team noticed that the number of supplies taken from the new cart was much higher than those taken from the old cart. Specifically, there were several items that nurses were not willing to go to another cart for, such as speculums or certain test kits, but were happy to grab if they were present. This implied that having carts stocked with these supplies meant that nurses would actually use them, providing more timely patient care.

Plans for and Barriers to Full Implementation

Due to resource constraints, the new drawers are not fully implemented throughout the department. Full implementation would require a large up-front cost: due to the reorganization of supplies, some supplies were moved to different drawers. Because hospital policy requires that all carts contain the same organization, it is not possible to gradually phase in new carts over an extended period. In addition, because of this reorganization, it is not possible to replace the cart drawers for all carts, 1 drawer at a time, since incomplete adoption would mean that some items would no longer be included. This is certainly a limitation of the team's design process—had considerations about gradual or modular adoption been considered earlier in the process, this would not have been an issue. However, the team is working to find the lowest-cost way to implement these new designs.

They are also working to mitigate unintended consequences arising from adjustments to the new cart system. Specifically, to improve the ease of adjustment, they have created a pamphlet showing an updated schematic of the cart with pictures of the contents of each full cart. A written list of the cart's contents and approximate capacity will also be provided to keep at each nursing station for training purposes and reference. They will label carts with the contents of each drawer (ie, drawer 1 will be labeled with IV supplies; drawer 2 will be labeled with wound care, diagnostics, and extra needles; drawer 3 will be labeled with miscellaneous small supplies; drawer 4 will be labeled with respiratory supplies; and drawer 5 will be labeled with flushes, urine cups, catheters, tubing), and the bottom of each section of the drawer will have a photo of the supply that belongs inside it. These improvements will facilitate stocking, assist with the transition, and make it easier to onboard new nurses. In addition, it will decrease the likelihood of drift over time, ensuring that staff use the partitions as directed.

Discussion

Principal Findings

This participatory design process highlighted the importance of co-design for improving workflow and supply management in an ED. By observing, interviewing, and directly working with the people who use the system the most, the team was better able to understand the key challenges that influenced the effectiveness and constrained the solution space. The final prototype increased the capacity for key items by up to 200%, enabling cart capacity to last 36 hours (3 shifts) rather than requiring restocking within the first 12-hour shift. In addition, similar to other studies [15,31], the reduction in response time finding supplies did not reach statistical significance, but trended toward a statistical result. This is in spite of the fact that the new design meant participants spent more time looking for objects that had been moved, a metric that would improve as users became more familiar with the new system. Notably, the improvement here is likely due to the new design reducing the number of trips to the clean supply room or searching for supplies from other carts, rather than from the more logical organization of the cart itself.

Overall, the findings of improved supply accessibility are in line with reports from other human factors–based redesigns of supply carts in other departments of the hospital [1,2,4,7,12-14]. While the constraints of this setting are not entirely analogous to other settings—here, the challenge was less about disorganization and more about a limited availability of stockers—they nevertheless suggest that cart organization, tailored to the needs of the staff and built with human factors principles in mind, plays a critical role in departmental workflows.

Even when the carts were fully stocked, nurses noted that the items that were present were there in the wrong quantities. Low-frequency items were present in large quantities, while there were hardly any of the most commonly used supplies. In fact, the supply manager told us that she had gone through the carts and had to toss hundreds of dollars' worth of unused supplies that had expired, an enormous waste of money and space for a department that had little to spare of either. This finding of overstocked items leading to wastage has been found in other sites as well [3,32] and suggests that periodic evaluation to ensure that carts can contain appropriate quantities of each supply is necessary. Our discovery of clusters of items frequently used together also matches the findings reported elsewhere [33].

In addition, the success of the quality improvement demonstrated the importance of participatory design, particularly the frequent consultation with stakeholders during the ideation and prototyping process, ensuring that the team's proposed solution was both effective and acceptable. Stakeholders were glad to be able to address a problem that mattered a great deal to them and helped to ensure that the new designs addressed their particular concerns about the old model. Co-design processes and design sprints are widely used in applied health research and are critical for ensuring that solutions focus on outcomes that are most important to end users [23,34,35].

Tailoring the cart's contents to the user workflow, and specifically minimizing the number of drawers that needed to be opened, improved user experiences. Although the cart's organization changed slightly, and users were less familiar with it, they still found it easier to use and preferred it over the standard cart. This is similar to prior findings that nurses preferred to use a cart that only required them to open 1 drawer per patient rather than a cart where everything was organized by tool but required multiple accesses [13].

In addition, nurses discussed how eager they were to be able to participate in the process. While this has been a persistent problem, engaging nurses in the redesign process gave them a voice, helping them feel like change was possible. This matches with previous research that found that nurse involvement in the co-design process improves nurse satisfaction [14].

The prototype created by this team is an example of how participatory design can create substantial improvement in a nursing workspace. The team was able to identify the key items nurses needed most and provide enough of those items to last multiple shifts without restocking, substantially improving workflow. The organization was designed according to nurses' priorities, focusing on ergonomics and efficiency using the same principles found in other studies [14,15]. While the solution could not solve the root causes of the issue—low funding and staffing shortages—it still made a meaningful, actionable contribution within the preexisting constraints by reducing the frequency of needed restocking. Once fully implemented, it is expected to save nurses a substantial amount of time per shift looking for items, reducing frustration and increasing the amount of time available for providing quality patient care. In addition, the durability of the acrylic, the work done to document and improve the ease of replicability of the system, and the buy-in from stakeholders all help to cement its durability and reduce the drift in restocking practices.

While the final design of the drawers in this project is customized to the specific hospital, the co-design process used demonstrates how stakeholder contribution and co-design can enable teams to make a meaningful impact in a low-resourced setting, working around structural barriers to improve nurse workflow and patient care.

Limitations

This project, while useful, was unable to address the root cause of the problem of supply shortages: the lack of a dedicated person who could simply restock all of the carts on a reliable basis. Despite this challenge, the implemented solution managed to ameliorate these effects by decreasing the time needed to restock and increasing the time the cart could go without restocking.

The project also encountered challenges regarding implementation—specifically, that because some items were moved to different drawers, and institutional policy required that all carts be organized the same way, implementation would need to occur all at once rather than being phased in gradually, requiring a larger up-front cost.

In addition, the data collection for the simulations was limited due to team availability and the availability of nursing staff to assist during their own busy shifts. The simulations also relied on convenience sampling: working with whatever staff were available at the time. Staff received the same prompts for both carts in the simulation; it is possible that thinking through the supplies beforehand made them faster at accessing the supplies the second time. We attempted to mitigate this response by randomizing which cart the team used first. Variable levels of experience with both the old and revised carts may also have led to bias in how quickly nurses found supplies in each cart [15]. Despite the lack of significance in time saved, the team was able to validate the use of the intervention through the increase in capacity for key supplies and through positive qualitative feedback from nursing staff indicating that, based on their past experiences, the cart would make a substantial improvement.

Finally, the designs for the cart interiors created during this project were tailor-made for this specific department and the cart system already in place, and would not necessarily generalize to other EDs. However, the processes used to design and validate the solution are readily transferable, and the designs could be tweaked to fit the unique contexts of other departments.

Conclusion

The co-design process used in this study demonstrates the importance of stakeholder engagement and contribution to quality improvement and the use of the double diamond approach in a health care setting. Stakeholder input was crucial for navigating around institutional barriers and ensuring that any proposed solutions would be accepted and implemented, rather than ignored. Indeed, the buy-in generated from participatory design ensured that the changes made would be sustainable. While the team’s designs were specific to the ED’s context, the process and insights can be applied to any other type of workplace improvement process. Specifically, we recommend engaging with stakeholders at every step of the way, particularly during the rapid prototyping phase. However, we suggest that future teams consider the ease of implementation, especially gradual implementation for budgetary reasons, while evaluating their designs.

Acknowledgments

The content is solely the responsibility of the authors and does not necessarily represent the official views of any funding agency. This work could not have been done without the incredible staff at the emergency department, whose insight and perspective were crucial. We greatly appreciate their time and willingness to speak with us. In addition, we are grateful to Helen Zhang for her perspectives on design.

Funding

KH is supported generally by the National Institutes of Health (5T32NS091006-10). This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. No external organization had any role in the data for this study.

Data Availability

Full anonymized data may be made available upon request.

Authors' Contributions

Conceptualization: AM, AW, KH, NR, SD, TH
Data curation: AM, AW, KH, SD, TH
Investigation: AM, AW, KH, SD, TH
Methodology: AM, AW, KH, SD, TH
Project administration: JD, NR, ON
Resources: AM, SD
Review & editing: AM, AW, KH, TH
Supervision: JD, NR, ON
Writing – original draft: AM, AW, KH, TH
Writing – review & editing: JD, NR, ON, SD

Conflicts of Interest

None declared.

Multimedia Appendix 1
Timeline of steps for cart intervention project.
[PNG File, 58 KB - humanfactors_v13i1e80861_app1.png]

Multimedia Appendix 2
List of interviewees for primary stakeholder interviews.
[DOCX File, 15 KB - humanfactors_v13i1e80861_app2.docx]

Multimedia Appendix 3
High-use items identified by clustering algorithm and nurse report.
[DOCX File, 18 KB - humanfactors_v13i1e80861_app3.docx]

Multimedia Appendix 4
Comparison of the number of items in the current nursing carts to the modified amounts in the prototype (complete inventory).
[DOCX File, 26 KB - humanfactors_v13i1e80861_app4.docx]

References

1. McLaughlin RC. Hospital extra: redesigning the crash cart: usability testing improves one facility's medication drawers. Am J Nurs 2003 Apr;103(4):64A-64F. [doi: 10.1097/00000446-200304000-00037]
2. Lefebvre MS, Milloy S, Joynt C. Optimal crash cart configuration for a surgical NICU: utilizing human factors principles. Adv Neonatal Care 2021 Aug 1;21(4):289-396. [doi: 10.1097/ANC.0000000000000814] [Medline: 33278105]
3. Silva VFD, Lazzari DD, Reisdorfer N, Michaelsen SC, Kuhnen AE. Analyzing the operational conditions of crash carts in clinical and surgical hospitalization units. Rev Esc Enferm USP 2021;55:e03693. [doi: 10.1590/S1980-220X2019040003693] [Medline: 33886918]
4. Cajigas I, Silva MA, Chang H, et al. EVD supply cart decreases time to ventriculostomy placement in the emergency department. Neurosurgery 2020;67(Supplement_1). [doi: 10.1093/neuros/nyaa447_301]
5. Fahrenkopf MP, Eichhorn MG. Development of a plastic surgery supply cart: patient outcomes and quality of care. Plast Reconstr Surg Glob Open 2019 Feb;7(2):e2111. [doi: 10.1097/GOX.0000000000002111] [Medline: 30881838]
6. Johnson K, Smith C, Barker LG, et al. Implementation of a resident-designed procedure cart in a busy emergency department. Cureus 2023 Jul;15(7):e42211. [doi: 10.7759/cureus.42211] [Medline: 37602021]
7. Maul E, Latham B, Westgate PM. Saving time under pressure: effectiveness of standardizing pediatric resuscitation carts. Hosp Pediatr 2016 Feb;6(2):67-71. [doi: 10.1542/hpeds.2015-0161] [Medline: 26794403]
8. Shultz J, Davies JM, Caird J, Chisholm S, Ruggles K, Puls R. Standardizing anesthesia medication drawers using human factors and quality assurance methods. Can J Anaesth 2010 May;57(5):490-499. [doi: 10.1007/s12630-010-9274-8] [Medline: 20143279]
9. Tucker AL, Spear SJ. Operational failures and interruptions in hospital nursing. Health Serv Res 2006 Jun;41(3 Pt 1):643-662. [doi: 10.1111/j.1475-6773.2006.00502.x] [Medline: 16704505]
10. Storfjell JL, Ohlson S, Omoike O, Fitzpatrick T, Wetasin K. Non-value-added time: the million dollar nursing opportunity. J Nurs Adm 2009 Jan;39(1):38-45. [doi: 10.1097/NNA.0b013e31818e9cd4] [Medline: 19104286]
11. Citolino CMF, Santos ES, Silva RDCG, Nogueira LDS. Factors affecting the quality of cardiopulmonary resuscitation in inpatient units: perception of nurses. Rev Esc Enferm USP 2015 Dec;49(6):908-914. [doi: 10.1590/S0080-623420150000600005] [Medline: 27419673]
12. Rousek JB, Hallbeck MS. Improving medication management through the redesign of the hospital code cart medication drawer. Hum Factors 2011 Dec;53(6):626-636. [doi: 10.1177/0018720811426427] [Medline: 22235525]
13. Agarwal S, Swanson S, Murphy A, Yaeger K, Sharek P, Halamek LP. Comparing the utility of a standard pediatric resuscitation cart with a pediatric resuscitation cart based on the Broselow tape: a randomized, controlled, crossover trial involving simulated resuscitation scenarios. Pediatrics 2005 Sep;116(3):e326-e333. [doi: 10.1542/peds.2005-0320] [Medline: 16061568]
14. Crimlisk JT, Doherty MM, Fernandes E, Leblanc E, Guarino R, Costello KV. Adult code cart redesign: clinical implications. Nursing (Auckl) 2018 Jul;48(7):58-61. [doi: 10.1097/01.NURSE.0000532760.94860.2e] [Medline: 29923923]
15. Frazier M, Webster K, Dewan M, et al. Improving usability of the pediatric code cart by combining lean and human factors principles. Pediatr Qual Saf 2023;8(4):e676. [doi: 10.1097/pq9.0000000000000676] [Medline: 37551262]
16. Probst CA, Carter M, Cadigan C, et al. Utilizing a human factors nursing worksystem improvement framework to increase nurses' time at the bedside and enhance safety. J Nurs Adm 2017;47(2):94-100. [doi: 10.1097/NNA.0000000000000446] [Medline: 28067682]
17. Tucker AL, Singer SJ, Hayes JE, Falwell A. Front-line staff perspectives on opportunities for improving the safety and efficiency of hospital work systems. Health Serv Res 2008 Oct;43(5 Pt 2):1807-1829. [doi: 10.1111/j.1475-6773.2008.00868.x] [Medline: 18522667]
18. Ford S. Nurses waste "an hour a shift" finding equipment. Nurs Times 2009;105(5):1. [Medline: 19266954]
19. Rochais É, Atkinson S, Bussières JF. Nursing perception of the impact of medication carts on patient safety and ergonomics in a teaching health care center. J Pharm Pract 2013 Apr;26(2):131-137. [doi: 10.1177/0897190012451908] [Medline: 22797833]
20. Lindsay MR, Lytle K. Implementing best practices to redesign workflow and optimize nursing documentation in the electronic health record. Appl Clin Inform 2022 May;13(3):711-719. [doi: 10.1055/a-1868-6431] [Medline: 35668677]
21. Graham-Dickerson P, Houser J, Thomas E, et al. The value of staff nurse involvement in decision making. J Nurs Adm 2013 May;43(5):286-292. [doi: 10.1097/NNA.0b013e31828eec15] [Medline: 23615371]
22. The Field Guide to Human-Centered Design, 1st edition: IDEO.org; 2015. URL: https://www.ideo.com/journal/design-kit-the-human-centered-design-toolkit [accessed 2026-01-08]
23. Peters S, Guccione L, Francis J, et al. Evaluation of research co-design in health: a systematic overview of reviews and development of a framework. Implement Sci 2024 Sep 11;19(1):63. [doi: 10.1186/s13012-024-01394-4] [Medline: 39261956]
24. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0—Standards for Quality Improvement Reporting Excellence—revised publication guidelines from a detailed consensus process. J Am Coll Surg 2016 Mar;222(3):317-323. [doi: 10.1016/j.jamcollsurg.2015.07.456] [Medline: 26385723]
25. Jamshed S. Qualitative research method-interviewing and observation. J Basic Clin Pharm 2014 Sep;5(4):87-88. [doi: 10.4103/0976-0105.141942] [Medline: 25316987]

26. Grady MP. Qualitative and Action Research: A Practitioner Handbook: Phi Delta Kappa Educational Foundation; 1998.
27. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
28. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. Int J Qual Methods 2023 Oct;22:WOS. [doi: 10.1177/16094069231205789]
29. Zenios SA, Makower J, Yock PG. Biodesign: The Process of Innovating Medical Technologies: Cambridge University Press; 2010. [doi: 10.1017/CBO9780511757853]
30. Onshape. URL: https://www.onshape.com/en/ [accessed 2026-01-17]
31. Chen DW, Chase VJ, Burkhardt ME, Agulto AZ. Using human factors design principles and industrial engineering methods to improve accuracy and speed of drug selection with medication trays. Jt Comm J Qual Patient Saf 2016;42(10):473-477. [doi: 10.1016/s1553-7250(16)42065-9] [Medline: 27712606]
32. Tsima BM, Rajeswaran L, Cox M. Assessment of cardiopulmonary resuscitation equipment in resuscitation trolleys in district hospitals in Botswana: a cross-sectional study. Afr J Prim Health Care Fam Med 2019 Oct 17;11(1):e1-e7. [doi: 10.4102/phcfm.v11i1.2029] [Medline: 31714118]
33. Hansoti B, Hahn E, Rao A, et al. Calibrating a chief complaint list for low resource settings: a methodologic case study. Int J Emerg Med 2021 May 19;14(1):32. [doi: 10.1186/s12245-021-00347-8] [Medline: 34011284]
34. Wu L, Zhu C, Seguin JP, et al. Designing culturally adapted digital mental health support tool for Chinese-speaking international students in Australia: a qualitative co-design study. JMIR Form Res 2025 Oct 21;9:e76695. [doi: 10.2196/76695] [Medline: 41186423]
35. Hopstaken JS, Koeneman M, Hooijer R, et al. How to improve pancreatic cancer network care using a human-centered design sprint. J Med Internet Res 2025 Sep 5;27:e55598. [doi: 10.2196/55598] [Medline: 40911870]

Abbreviations

ED: emergency department
IV: intravenous
SQUIRE: Standards for Quality Improvement Reporting Excellence

---

Edited by E Borycki; submitted 07.Aug.2025; peer-reviewed by EP Larsen, K Webster; revised version received 10.Nov.2025; accepted 13.Nov.2025; published 28.Jan.2026.

Please cite as:
Hefter K, Dreibelbis S, Haupt T, McIver A, Wang A, Dwight J, Nwodim O, Ray N
Redesign of Bedside Supply Carts to Improve Emergency Department Workflows: Mixed Methods Participatory Design
JMIR Hum Factors 2026;13:e80861
URL: https://humanfactors.jmir.org/2026/1/e80861
doi:10.2196/80861

Wearable Device Photoplethysmography As a Viable Tool to Longitudinally Monitor Vasoconstriction Biomarkers for Predicting Vaso-Occlusive Crisis in Sickle Cell Disease: Feasibility and Validation Study

Payal Shah[1], MS; Sabrina Sy[2], BS; Mingjing Chen[2], BS; Michael CK Khoo[2], PhD; Thomas D Coates[1], MD; Saranya Veluswamy[1], MD

[1]Division of Hematology, Cancer and Blood Disease Institute, Children's Hospital Los Angeles, University of Southern California, 4650 Sunset Boulevard, Los Angeles, CA, United States

[2]Alfred E. Mann Department of Biomedical Engineering, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Saranya Veluswamy, MD
Division of Hematology, Cancer and Blood Disease Institute, Children's Hospital Los Angeles, University of Southern California, 4650 Sunset Boulevard, Los Angeles, CA, United States

Abstract

Background: Entrapment of sickled red blood cells in the microvasculature leads to sudden painful vaso-occlusive crises (VOCs) in sickle cell disease (SCD). This is potentially triggered by autonomic nervous system–mediated vasoconstriction in the microvasculature. Indeed, vasoconstriction biomarkers derived from a single night of laboratory-based fingertip photoplethysmography (PPG) recording were predictive of a higher frequency of future VOC in SCD. Noninvasive, remote, and longitudinal monitoring of autonomic vasoreactivity will facilitate the development of predictive biomarkers of imminent VOC.

Objective: This study aimed to assess the feasibility and performance of a wearable wristband device to longitudinally monitor nocturnal peripheral autonomic vasoreactivity and to cross-validate the vasoconstriction parameters across the “gold-standard” finger sensor.

Methods: A total of 12 patients with SCD and 6 healthy controls were recruited to wear a wristband device (Biostrap) with a PPG sensor on a nightly basis. For cross-validation studies, 50% (3/6) controls wore both the wristband and a sleep monitoring device (AliceNightOne) with a finger PPG sensor. We quantified autonomic vasoreactivity by processing PPG signals and deriving vasoconstriction parameters—magnitude of vasoconstriction (Mvasoc) and photoplethysmography amplitude coefficient of variation (PPGampCV). We performed a correlation analysis of the vasoconstriction parameters within each device to investigate whether Mvasoc and PPGampCV can be used as surrogate markers of vasoconstriction, and then cross-validated the PPGampCV across the wristband and finger PPG devices.

Results: A total of 131 nocturnal PPG recordings were made with a wristband device (1 - 19 nights per participant; patients with SCD: n=79, 60%; controls: n=52, 40%). A total of 9 nocturnal recordings (3 nights per participant) were made with both wristband and finger sensor devices. Longitudinal continuous PPG recordings were feasible with the wearable device, with significant within-night and night-to-night variability in vasoconstriction parameters, suggesting dynamic changes in autonomic vasoreactivity. Mvasoc and PPGampCV significantly correlated within devices—the maximum overnight correlation was 0.82 (*P*<.001) for the finger sensor and 0.69 (*P*<.001) for the wristband sensor, suggesting that PPGampCV can serve as a surrogate for Mvasoc. Cross-validation analysis of PPGampCV across wristband and fingertip sensors showed statistically significant correlations on all 9 nights (overnight correlation coefficient ranging from 0.24 - 0.7), with some nightly segments of PPGampCV showing very strong correlation across devices.

Conclusions: Wearable wristband devices are feasible tools for the collection of continuous PPG measurements and vasoconstriction parameters, which serve as objective markers of autonomic vasoreactivity in users with and without SCD. We have optimized the methods of quantifying vasoconstriction from wearable device PPG signals, and cross-validated them with standardized sensors. These findings enable large-scale, real-time monitoring of autonomic vasoreactivity along with pain outcomes for the development of vasoconstriction parameters as biomarkers imminent VOC in patients with SCD. This biomarker also has the potential to impact other diseases involving autonomic vascular dysregulation.

(JMIR Hum Factors 2026;13:e75465) doi:10.2196/75465

KEYWORDS

sickle cell disease; wearable devices; microvascular blood flow; vasoconstriction; remote monitoring

Introduction

Vaso-occlusive crisis (VOC) due to obstruction of microvascular blood flow by sickled red blood cells is the prime reason for the morbidity and mortality in sickle cell disease (SCD) [1]. Peripheral vasoconstriction in the microvasculature due to dysregulated autonomic nervous system (ANS) responses and the subsequent reduced microvascular blood flow increases red blood cell transit time in the microvasculature, thus potentiating VOC [2-7]. Our recent work on the neurovascular physiology of SCD in humans has established that ANS-mediated decreases in microvascular blood flow, quantified by a novel vasoconstriction biomarker, strongly predict future VOC [8]. This novel vasoconstriction index, derived from the amplitude of finger photoplethysmography (PPG) signals, is a composite measure of the dynamic changes in peripheral blood flow occurring over time and is reflective of sympathetic neural inputs to the microvasculature. A single nocturnal measurement of this index in patients with SCD showed that a higher magnitude of vasoconstriction (Mvasoc) was predictive of increased hospitalizations for VOC in the upcoming years. However, day-to-day variations in autonomic vasoreactivity and their implications for imminent VOC risk remain unknown. All our previous measurements of peripheral vasoreactivity and ANS activity were performed in a well-controlled laboratory environment using standard finger PPG sensors to measure microvascular blood flow. Longitudinal remote monitoring of vasoconstriction events and ANS activity has a high potential for predicting and monitoring individualized pain episodes and is the necessary next step to develop nocturnal vasoconstriction indices as a predictor of imminent VOC.

Current wearable devices use optical PPG sensors to monitor heart rate variability and oxygen saturation. However, the raw PPG signals are often inaccessible before they go through automatic gain and processing, such that key information from the amplitude of the raw PPG signal is lost. Moreover, in the most commonly used commercial wearable devices, the PPG signal is measured from the wrist [9] or at the base of a finger [10,11]. These anatomic locations are not as highly perfused as the fingertip, and the PPG waveforms may be expected to be morphologically different from the fingertip due to differences in local vascular physiology and lower signal-to-noise ratios. On the other hand, wearable device PPG measurements are likely more reliable based on user compliance and practical considerations.

In this study, we established the use of PPG signals from a wrist-worn wearable device for real-time, remote monitoring of autonomic vasoreactivity and measurement of vasoconstriction events during sleep. The Biostrap EVO is a commercially available wrist-worn device (Biostrap LLC) that uses an optical PPG sensor to record standard continuous physiological data [12], including heart rate, computed respiratory activity, oxygen saturation (SpO2), and sleep and activity patterns [13]. Uniquely, our team was also able to access the raw PPG waveforms from the device, which is critical to quantify our key vasoconstriction parameters. The AliceNightOne (ANO) is a Food and Drug Administration–approved home sleep testing device (Koninklijke Philips N.V.), primarily designed to detect sleep apnea [14]. The 3-sensor device with a nasal cannula, finger PPG sensor, and chest belt detects nasal airflow pressure, SpO2, heart rate, respiratory effort, and body positioning. The ANO uses medical-grade “gold-standard” finger PPG probes to measure heart rate and oxygenation. The aims of this study were to

1. Assess the feasibility of wristband wearable devices to remotely monitor nocturnal peripheral autonomic vasoreactivity
2. Optimize the measurement of vasoconstriction parameters from wearable device PPG in patients with SCD and healthy controls.
3. Cross-validate the vasoconstriction parameters derived from a wearable wristband sensor (Biostrap) and a “gold-standard” finger sensor (ANO).

Methods

Ethical Considerations

The study was approved by the Institutional Review Board of the Children’s Hospital, Los Angeles, and all participants provided informed consent or assent to participate in the study. The study was conducted in accordance with the Declaration of Helsinki and the data was de-identified as per the Health Insurance Portability and Accountability Act (HIPAA), and subjects were compensated for their participation.

Study Setup

Patients with SCD and healthy controls underwent autonomic monitoring with a wristband sensor to measure continuous PPG during sleep. Recordings were automatically triggered at sleep onset and paused for 5 minutes every hour to allow syncing of data wirelessly to a phone-based app. The raw PPG waveforms were made accessible from a cloud-based server, exported as .csv files, and selected for analyses based on the quality and completeness of the overnight recording. Raw signals were first screened to determine if there were any segments with significant motion artifacts and were compared alongside accelerometer data also available from the Biostrap cloud server.

For the cross-validation of vasoconstriction parameters, a subset of the control participants wore the ANO sleep device with a finger sensor simultaneously with the Biostrap wristband sensor. Sleepware G3 (version 3.9.7.0; Respironics Inc) provides the capability of storing acquisitions from an ANO sleep device on a cloud-based server and exporting the .edf files (European Data Format) [15]. We converted data from both wristband and finger sensors into .Mat files (MATLAB) to compare them synchronously on 1 platform and derive the physiological parameters of interest.

PPG Signal Processing and Derivation of Vasoconstriction Parameters

Signal Processing

The peak-to-trough amplitude of each pulse in the raw PPG signal reflects the pulsatile change in arteriolar blood volume, with a lower amplitude indicating vasoconstriction (Figure 1). The photoplethysmography amplitude (PPGamp), along with the pulse duration, was calculated from both wrist and finger raw PPG signals and divided into consecutive 15-minute segments (Figure 2A). First, segments with clear signal loss were visually identified and manually trimmed or eliminated. Second, outliers were removed by replacing values exceeding the 98th percentile within each segment with the 98th percentile value. Finally, the signal was normalized to its 95th percentile. After linear detrending, the respiratory influence (frequency >0.15 Hz) was filtered out from both wristband and finger PPGamp signals (Figure 2B). These processed “clean” PPGamp signals from each device were used to calculate corresponding vasoconstriction parameters.

Since each device had its own independent time base, the corresponding finger and wristband PPGamp signals had to be aligned in time in order to facilitate comparison of vasoconstriction parameters across devices. The cross-correlation function is a standard method in time series analysis that is used to assess the degree of linear similarity between 2 signals and has been comprehensively described in the literature [16]. We used this technique to compare and align the PPGamp signals from the 2 different devices. The PPGamp time series from each device was first autocorrelated to ensure that these signals represent physiological measurements rather than random noise. Thereafter, corresponding segments of the finger and wristband PPGamp signals underwent multiple time shifts over the other to locate the point of maximum positive cross-correlation within a +30 or –30 second window, thus synchronizing the PPGamp signals from both devices (Figure 2C).

Figure 1. Photoplethysmography (PPG) amplitude calculated from peak-to-trough of each pulse of raw PPG signal with a decrease in amplitude signifying vasoconstriction.

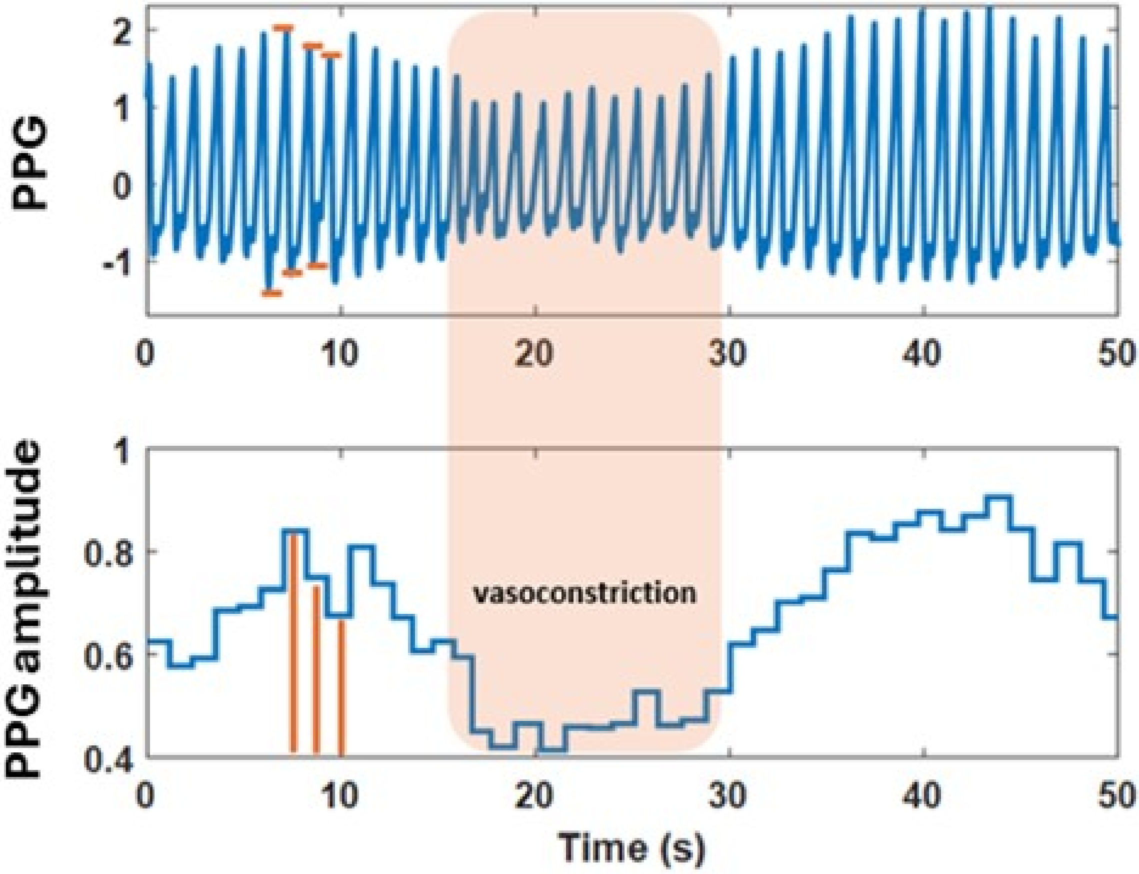
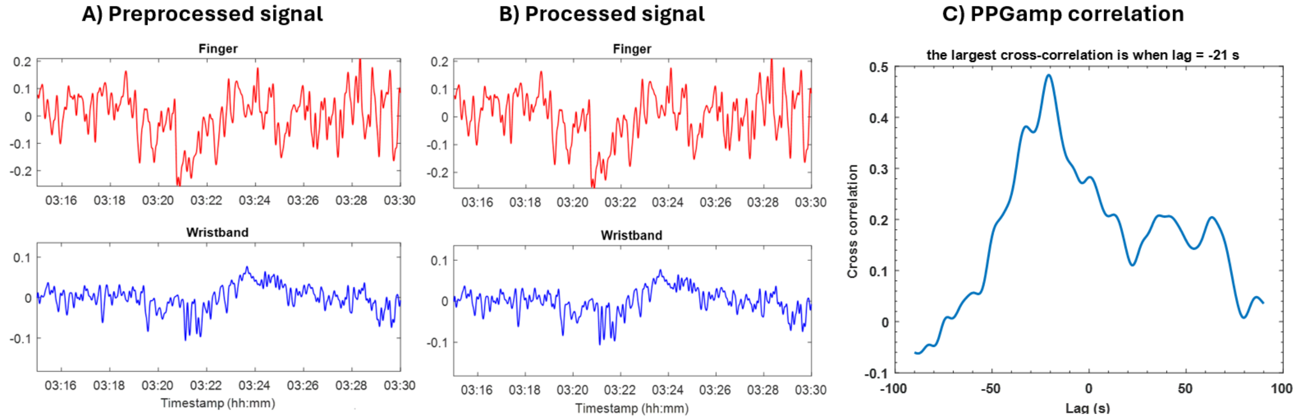


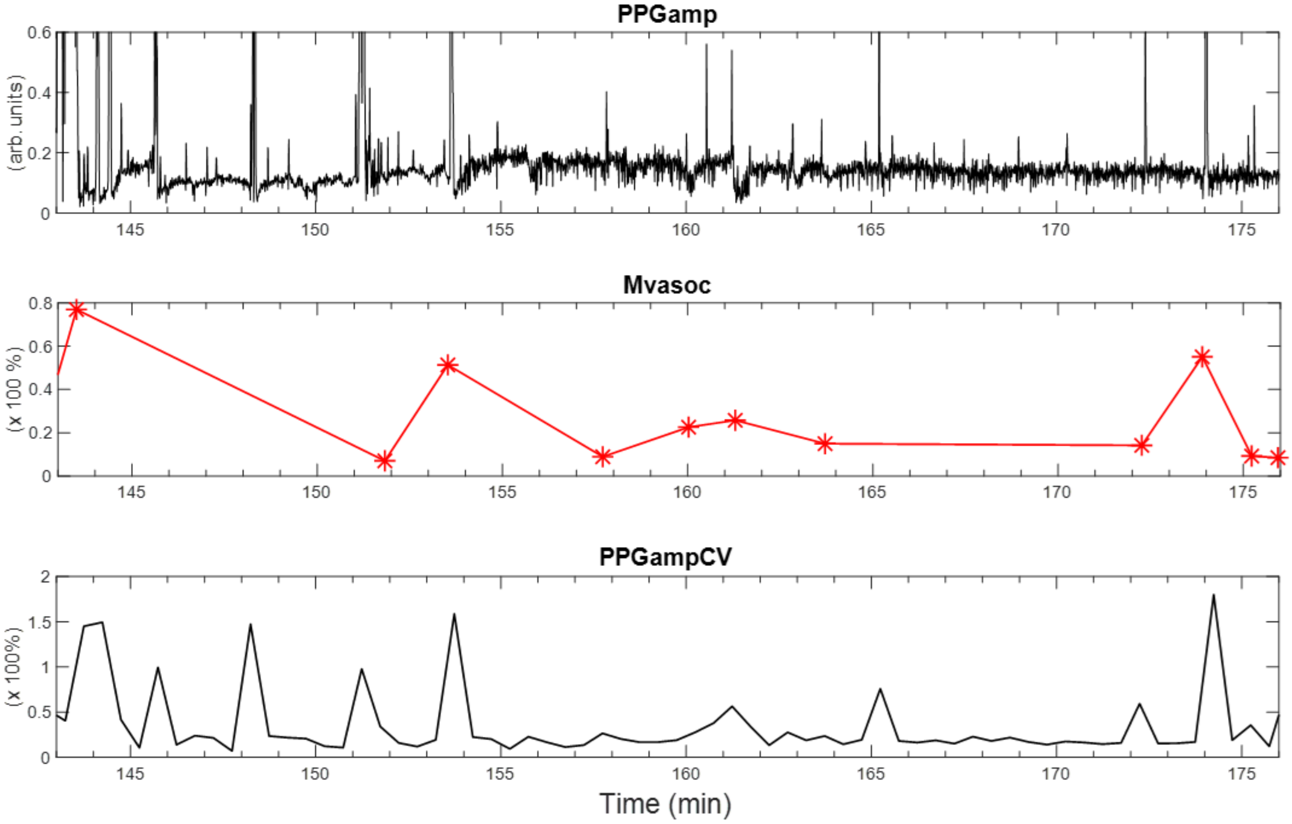
Figure 2. Representative 15-minute PPGamp segments from finger and wristband PPG sensors. (A) Preprocessed signals depicting finger and wristband PPGamp extracted from raw PPG signals. (B) Processed signals depicting “clean” finger and wristband PPGamp after removal of artifacts and respiratory influence. (C) Cross-correlation of the finger and wristband PPGamp segments. In this example, the maximal correlation occurs at a lag of −21 seconds. PPG: photoplethysmography; PPGamp: photoplethysmography amplitude.



Calculation of Vasoconstriction Parameters: Mvasoc and PPGampCV

We have previously introduced Mvasoc as a unique nocturnal vasoconstriction biomarker that takes into account the frequency, magnitude, and duration of autonomic-mediated spontaneous vasoconstrictions occurring each night [8]. The algorithm was developed to detect and quantify a vasoconstriction event as a significant reduction in PPGamp below the preceding 15-second baseline. Mvasoc is subsequently defined as the ratio of the area of the drop below baseline in PPGamp and the duration of the vasoconstriction, normalized by the baseline PPGamp level (described in detail in the supplement of Chalacheva et al [8]). Because many vasoconstriction events can occur during the night, the median Mvasoc value per night is used as a quantitative index of the user’s nocturnal vasoconstriction behavior during a given night [8]. However, when the Mvasoc algorithm is applied to PPGamp signals, the rule-based measure can be sensitive to time-varying changes in signal-to-noise ratio, and thus the time locations and amplitude of the detected vasoconstrictions can vary across PPG devices. To overcome this limitation, we used a second measure to quantify the dynamic variability in beat-to-beat PPGamp. More specifically, the coefficient of variation (CV; CV=SD/mean) of PPGamp within a sliding time window of 5 minutes was calculated every 30 seconds. In initial studies, this parameter (ie, PPGampCV) correlated closely and robustly with Mvasoc, while simultaneously eliminating the need for a rule-based algorithm to detect “significant vasoconstriction” events. Figure 3 is a representation of Mvasoc and PPGampCV parameters from a wristband PPGamp segment.

Figure 3. The first row represents a PPGamp segment from wristband sensor, and the next 2 rows show corresponding Mvasoc and PPGampCV parameters derived from the PPGamp. PPGamp: photoplethysmography amplitude; PPGampCV: photoplethysmography amplitude coefficient of variation; Mvasoc: magnitude of vasoconstriction.



Within-Device Correlation of Mvasoc and PPGampCV Vasoconstriction Parameters

To rigorously assess whether PPGampCV can be used as a surrogate for Mvasoc, we performed a correlation analysis of the 2 parameters as measured by the fingertip sensor in the ANO device and wrist sensor in the Biostrap device and derived the following within-device correlation coefficient (r_{wd}) values:

1. Overnight r_{wd}: represents the correlation between vasoconstriction parameters Mvasoc and PPGampCV for the whole night
2. Peak r_{wd}: represents the segment with the maximum correlation of Mvasoc and PPGampCV out of all the 15-minute segments on a given night

Across-Device Cross-Validation of Vasoconstriction Parameter

To determine if the wrist device gave the same results as the fingertip sensor, we correlated the PPGampCV parameter from the wristband and “gold-standard” finger sensors worn at the same time. The corresponding 15-minute segments from both devices were only considered for cross-correlation if at least 3 vasoconstriction events were detected within each segment. We correlated the PPGampCV between the finger and wrist PPGamp each night and derived the following across-device correlation coefficient (r_{ad}) values:

1. Overnight r_{ad}: represents the correlation of PPGampCV across wrist and finger sensor for all 15-minute segments per night
2. Peak r_{ad}: represents the segment with the maximal correlation of PPGampCV out of all corresponding 15-minute segments per night.

In all the analyses mentioned earlier, *P* values of <.05 suggested that the correlation between signals was statistically significant.

Results

Longitudinal Monitoring of Vasoconstriction Parameters With Wristband Device

A total of 12 patients with SCD and 6 control participants participated in the study. The mean age of the sample was 25 (range 14 - 40) years. The mean hemoglobin was 8.6 g/dl (SD 1.9 g/dl) and 14.1 g/dl (SD 1.8 g/dl) for patients with SCD and controls, respectively. A total of 10 (83%) of the 12 patients with SCD reported the use of disease-modifying therapy with hydroxyurea at the time of the study. A total of 131 nocturnal PPG recordings were made with wristband devices in all participants. Patients with SCD had a total of 79 (60%) nocturnal recordings (range 1 - 15 nights), and controls had 52 (40%) nocturnal recordings (range 1 - 19 nights). While participants wore the wristband for a variable number of nights due to device availability, there were no reported logistical issues with functioning of the device, and participants reported ease of use.

Longitudinal measurements of Mvasoc and PPGampCV vasoconstriction parameters showed significant variability within each night and from night to night. A representative participant’s longitudinal data are depicted in Figure 4 and show that there is a wider range of vasoconstriction activity on nights 10 and

11, with an increase in the median Mvasoc and PPGampCV noted on those nights, reflecting increased autonomic vasomotor activity. On visual inspection, the night-to-night changes in Mvasoc and PPGampCV appeared to closely track each other, suggesting a potential correlation between the 2 parameters. Indeed, the median Mvasoc of each night was closely correlated with the corresponding median PPGampCV both for patients with SCD and controls in all the nightly data (patients with SCD: R^2=0.79; controls: R^2=0.77; Figure 5).

Figure 4. Longitudinal monitoring of vasoconstriction parameters Mvasoc and PPGampCV in a patient with sickle cell disease from the wristband device. The box plot shows the median and IQR per night, and the midline shows the median of these parameters across all nights. PPGampCV: photoplethysmography amplitude coefficient of variation; Mvasoc: magnitude of vasoconstriction.

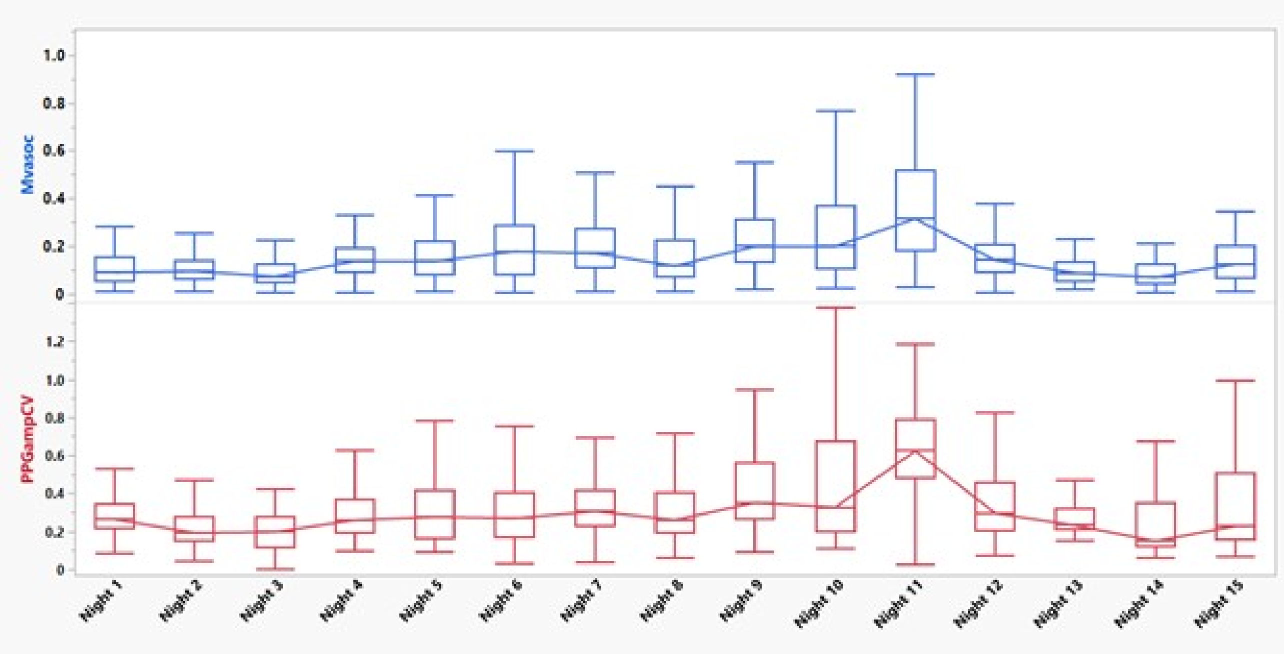
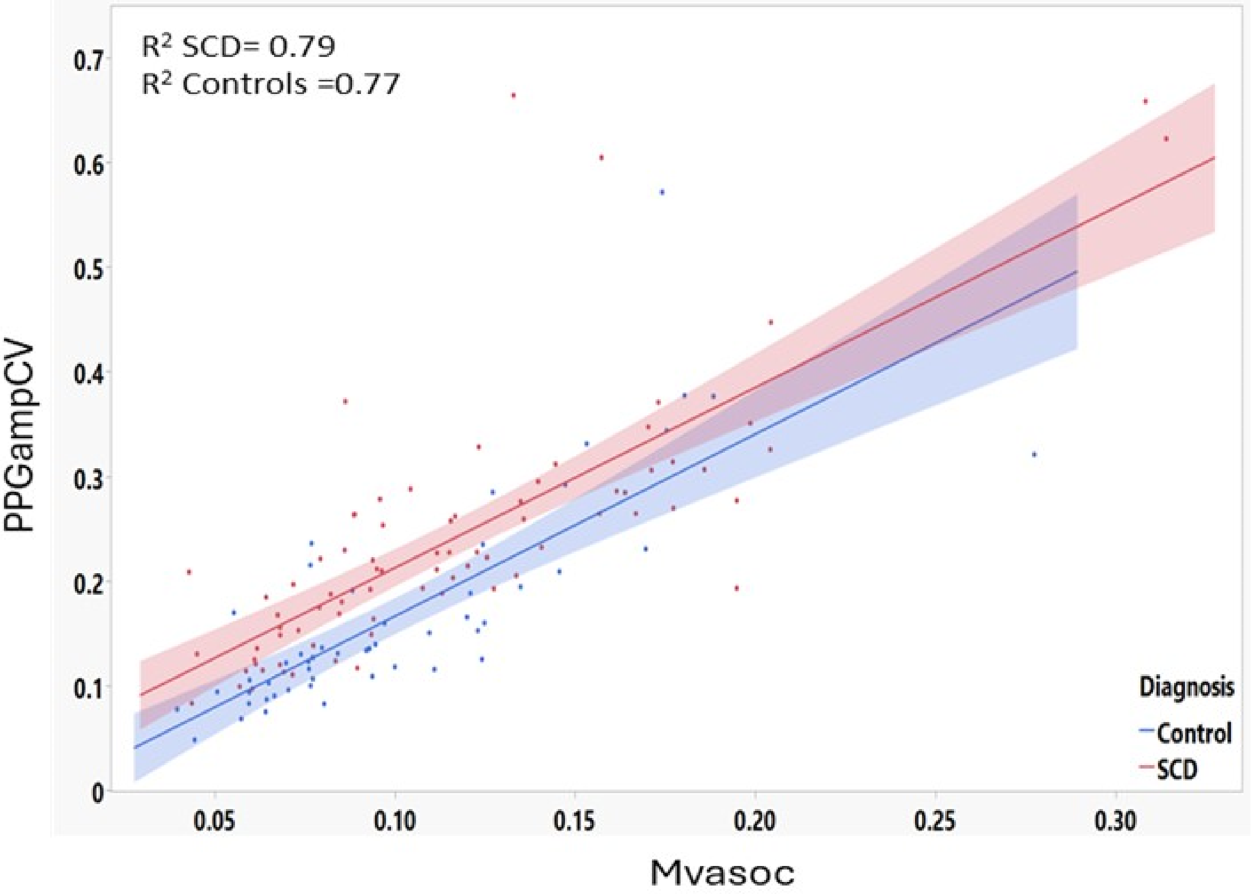


Figure 5. Correlation of median Mvasoc with corresponding median PPGampCV for all recorded nights from the wristband device among patients with SCD and controls. Mvasoc: magnitude of vasoconstriction; PPGampCV: photoplethysmography amplitude coefficient of variation; SCD: sickle cell disease.



Within-Device Correlation of Vasoconstriction Parameters (Mvasoc and PPGampCV)

While median Mvasoc and PPGampCV track closely in longitudinal nightly recordings, we further analyzed the correlation of these parameters during time segments within individual nightly recordings. Three control participants wore both finger and wristband PPG devices for a total of 9 nights, and each nighttime recording was divided into 15-minute segments. Table 1 shows the number of vasoconstriction (Mvasoc) events detected each night from both the finger sensor and wristband devices. There were a total of 113 viable segments over 9 nights, with an average of 10 and 11 vasoconstriction events detected during each segment from the finger and wristband sensors, respectively.

Table . Number of vasoconstriction events (magnitude of vasoconstriction) detected each night from the finger and wristband sensors.

	Finger sensor, n	Wristband, n
Participant 1		
Night 1 (5 segments)	34	56
Night 2 (8 segments)	90	104
Night 3 (12 segments)	140	152
Participant 2		
Night 1 (20 segments)	186	226
Night 2 (18 segments)	175	195
Night 3 (8 segments)	71	120
Participant 3		
Night 1 (19 segments)	206	214
Night 2 (18 segments)	200	190
Night 3 (5 segments)	40	57

Cross-correlation of Mvasoc and corresponding PPGampCV values were calculated for each segment of these 9 nights. Overnight r_{wd} correlation and segmental peak r_{wd} correlation coefficient data for all 9 nights (Table 2) show that Mvasoc was significantly correlated with PPGampCV each night within both devices and within almost all segments. The highest overnight r_{wd} correlation was 0.82 ($P<.001$) for the finger sensor and 0.69 ($P<.001$) for the wristband for all nights. The highest peak r_{wd} correlation was 0.99 ($P<.05$) in both devices for all nights (113 segments). Figure 6 shows the within-device overnight r_{wd} correlation between Mvasoc and PPGampCV for a representative night.

Table . Cross-correlation between magnitude of vasoconstriction (Mvasoc) and photoplethysmography amplitude coefficient of variation (PPGampCV) using overnight correlation coefficient (overnight r_{wd}) and peak correlation coefficient (peak r_{wd}) within wristband (Biostrap) and finger sensors (AliceNightOne).

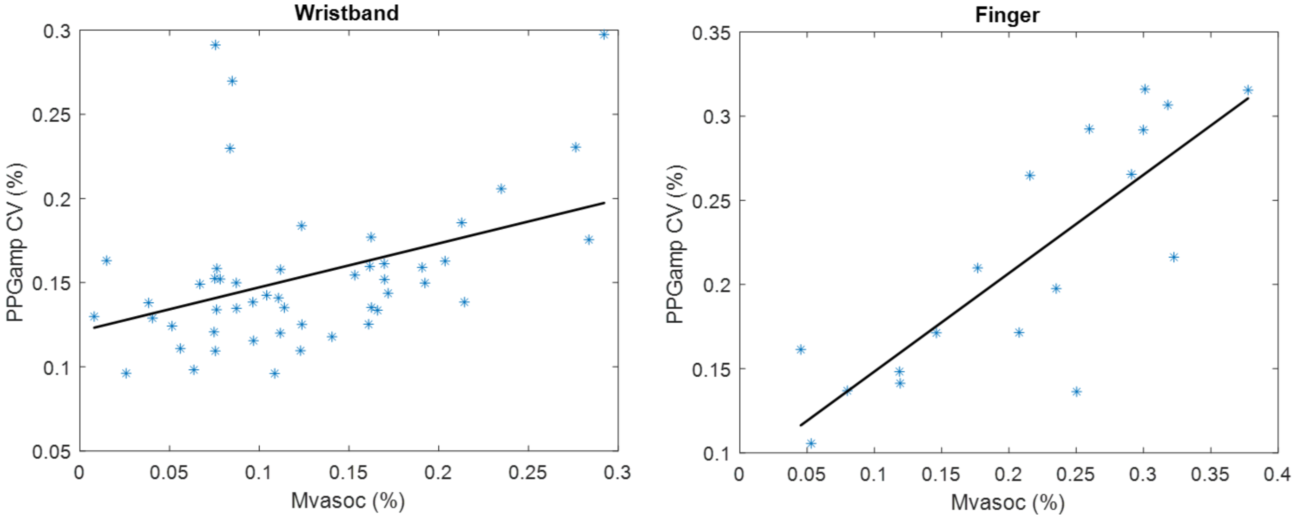
Mvasoc versus PPGampCV correlation	Finger sensor	Wristband
Participant 1		
Night 1 (5 segments)		
Overnight r_{wd}[a]	0.82	0.40
Peak r_{wd}[b]	0.90	0.74
Night 2 (8 segments)		
Overnight r_{wd}	0.35	0.47
Peak r_{wd}	0.93	0.78
Night 3 (12 segments)		
Overnight r_{wd}	0.51	0.41
Peak r_{wd}	0.82	0.80
Participant 2		
Night 1 (20 segments)		
Overnight r_{wd}	0.32	0.69
Peak r_{wd}	0.91	0.96
Night 2 (18 segments)		
Overnight r_{wd}	0.29	0.48
Peak r_{wd}	0.97	0.93
Night 3 (8 segments)		
Overnight r_{wd}	0.47	0.34
Peak r_{wd}	0.99	0.92
Participant 3		
Night 1 (19 segments)		
Overnight r_{wd}	0.40	0.49
Peak r_{wd}	0.80	0.99
Night 2 (18 segments)		
Overnight r_{wd}	0.37	0.51
Peak r_{wd}	0.79	0.95
Night 3 (5 segments)		
Overnight r_{wd}	0.33	0.30
Peak r_{wd}	0.95	0.50[c]

[a] Overnight r_{wd}**:** correlation coefficient value of vasoconstriction parameters, Mvasoc and PPGampCV, for the whole night; $P<.05$.

[b] Peak r_{wd}**:** correlation coefficient value of segment with highest correlation out of all 15-minute segments per night; $P<.05$.

[c] $P>.05$.

Figure 6. Example of overnight correlation (r_{wd}) between vasoconstriction parameters, Mvasoc and PPGampCV, for 1 control 1 night in wristband (Biostrap) and finger sensor (AliceNightOne). Panel (a wristband; *r*=0.40; *P*<.005) and AliceNightOne device in panel (b finger sensor; *r*=0.82; *P*<.0001). The significant correlation between Mvasoc and PPGampCV assured us that PPGampCV can be used as a surrogate marker of the magnitude of vasoconstriction. Mvasoc: magnitude of vasoconstriction; PPGampCV: photoplethysmography amplitude coefficient of variation.



Across-Device Cross-Validation of Vasoconstriction Parameters (CV)

Since the PPGampCV serves as a surrogate marker of vasoconstriction and is overall easier to derive from the PPG and compare across devices, we chose the PPGampCV to perform cross-validation of vasoconstriction parameters between the 2 devices. The highest overnight r_{ad} cross-correlation for all nights was 0.70, and the highest peak r_{ad} cross-correlation for all nights was 0.99. Table 3 shows the PPGampCV overnight correlation (r_{ad}) coefficient and segmental peak (r_{ad}) correlation coefficient data for all 9 nights (123 segments). Figure 7 shows an example of correlation analysis of PPGampCV across devices in 1 corresponding 15-minute segment.

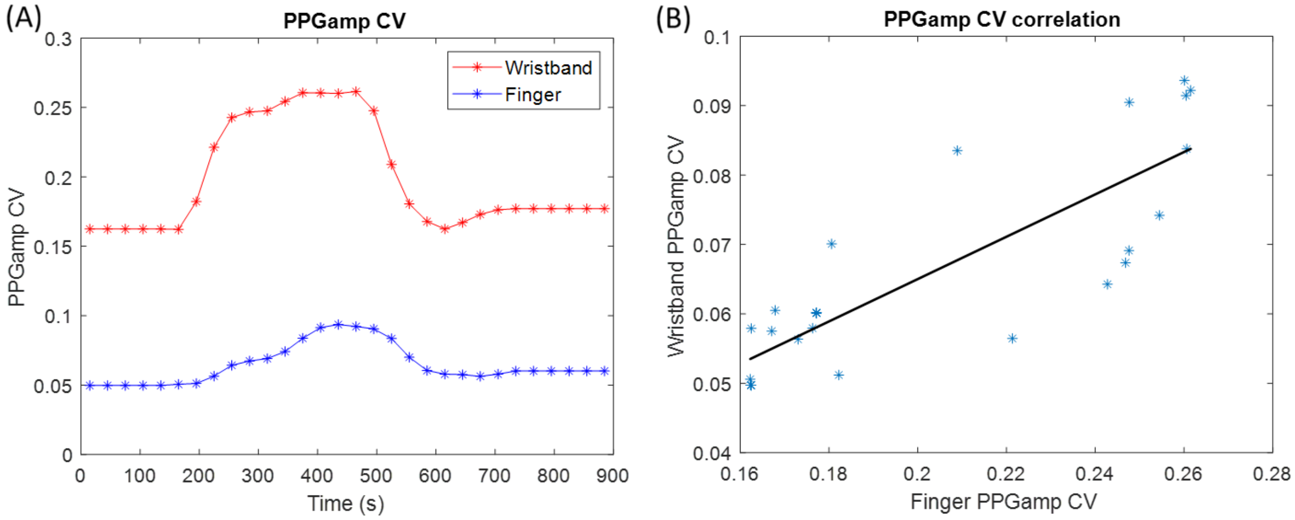
Table . Overnight r_{ad} correlation and peak r_{ad} correlation coefficients of photoplethysmography amplitude coefficient of variation between wristband (Biostrap) and finger sensor (AliceNightOne).

	Overnight r_{ad}[a]	Peak r_{ad}[b]
Participant 1		
Night 1 (5 segments)	0.24	0.76
Night 2 (8 segments)	0.50	0.90
Night 3 (12 segments)	0.70	0.89
Participant 2		
Night 1 (20 segments)	0.49	0.97
Night 2 (18 segments)	0.57	0.99
Night 3 (8 segments)	0.52	0.96
Participant 3		
Night 1 (23 segments)	0.36	0.97
Night 2 (24 segments)	0.62	0.96
Night 3 (5 segments)	0.65	0.56

[a] Overnight r_{ad}: represents the correlation coefficient value for all 15-minute segments per night; *P*<.05.

[b] Peak r_{ad}: represents the correlation coefficient value of segment with highest correlation out of all 15-minute segments per night; *P*<.05.

Figure 7. Correlation of PPGampCV between the wristband and finger device for a representative 15-minute segment of 1 night. The left panel (A) shows the corresponding PPGampCV values from both devices over time, and the right panel (B) shows the correlation of PPGampCV values between the 2 devices (*r*=0.84; *P*<.001). PPGampCV: photoplethysmography amplitude coefficient of variation.



Discussion

Principal Findings

The sudden, unpredictable onset of VOC remains the major cause of morbidity and the leading cause of hospitalizations in patients with SCD [17]. While some patients experience an aura or premonition that a VOC is about to occur [18], there are no objective biomarkers that can predict the risk of an imminent VOC and thus open up a therapeutic window to abort its propagation. We have introduced nocturnal vasoconstriction indices of autonomic peripheral vasoreactivity that have the potential to be predictive biomarkers of impending VOC crisis. This study establishes the feasibility of measuring these biomarkers in a real-world setting with a wearable device, such that the vasoconstriction biomarker can be validated in larger populations. To the best of our knowledge, this is the first study that uses the PPG signal from a wearable-based sensor system to quantify microvascular blood flow changes accompanying peripheral vasoconstriction.

We successfully used the Biostrap wearable wristband device to longitudinally monitor raw PPG signals and derive autonomic vasoconstriction parameters. Participants reported ease of use, and no significant barriers were identified in using the device. There was substantial night-to-night variability in the median Mvasoc and PPGampCV parameters within participants, which implies varying levels of sympathetic vasoreactivity. This has potential implications for tracking pain and imminent VOC risk, as SCD physiology dictates that vasoconstriction and resultant decrease in microvascular perfusion increase the likelihood of vaso-occlusion.

Although Mvasoc is a direct quantification of vasoconstriction events from the PPGamp signal, we found that this parameter is more prone to artifacts that can plague wearable device signals. The PPGampCV is a more robust measure of variability of the PPGamp amplitude that closely reflects Mvasoc and is an acceptable surrogate for monitoring vasoconstriction. The ANO with a fingertip PPG sensor is a widely used Food and Drug Administration–approved home sleep device that represents a reduced version of the clinical polysomnography. The vasoconstrictions measured at the fingertip are generally stronger than the corresponding events measured at the wrist, and thus the signal-to-noise ratio of the fingertip signal is higher than that of the wrist signal. Physiologically, this is to be expected since the fingers are highly innervated by alpha-adrenergic fibers and are also highly vascularized along with large numbers of arteriovenous anastomoses, compared to the wrist. This results in much faster and stronger vasoconstrictive responses in the fingers vis-à-vis the wrist. These differences in perfusion also have the potential to impact the cross-correlation values of the PPGampCV signals between the fingertip and the wrist. Using a cutoff of ≥0.5 as an acceptable correlation coefficient, we still see that there is mostly moderate to very strong correlation of PPGampCV between ANO and Biostrap. This shows that the wristband performs comparably to the “gold-standard” fingertip PPG while being much more practical to use. The overall correlation of PPGampCV for each night also remained statistically significant and acceptable between devices, suggesting that the wristband device remains a viable and feasible tool to deploy for longer-term, easy-to-use remote longitudinal monitoring of autonomic vasoreactivity.

This study is also unique in the sense that most wearable devices that purport to provide information about autonomic function focus on heart rate variability and SpO2 changes, but do not monitor changes in PPG amplitude, indicative of peripheral vasoconstriction [9]. Our Mvasoc algorithm and PPGampCV biomarker can be adapted to any wearable device that uses the PPG signal to measure ANS activity and quantify vasoconstriction in users with and without SCD. Promising preliminary data suggest a correlation between the Mvasoc parameter and the intensity of daily pain in patients with SCD. In 10 patients with SCD from this same cohort, we successfully tracked daily pain data using a SMS text message–based red cap survey, along with wristband recordings. Presence of pain and intensity of pain were tracked on a scale of 1 to 10. Of the 68 nocturnal wristband recordings with concurrent pain diary entries, 25 days of pain were reported with a pain intensity ranging from 3 to 8. While the data were not powered toward the prediction of a VOC crisis given the limited number of nights of autonomic vasoreactivity monitoring, we found that a higher vasoconstriction index (Mvasoc) and a greater parasympathetic withdrawal in the heart rate variability patterns predicted greater intensity of pain the next day on multivariate regression analysis [19]. While these data hint at possible temporal associations between autonomic vasoreactivity patterns and pain, they need to be validated with longer-term longitudinal monitoring of vasoconstriction parameters in larger patient populations with SCD in order to develop it as a predictive tool for imminent VOC.

In conclusion, our pilot data show that the implementation of PPG from wearable sensors has a high potential for the development of an autonomic vasoreactivity biomarker for VOC in SCD. The portability, cost-effectiveness, and low maintenance of these wearable sensors can make them suitable for remote use. Moreover, these vasoconstriction parameters derived from PPG amplitude can serve as easily accessible and readily quantifiable biomarkers of tissue perfusion that could be valuable in other disease states, such as migraine and coronary artery disease, where vasoreactivity plays a role in disease pathology.

Limitations

The small sample size of nocturnal recordings, especially with the ANO device, makes it difficult to generalize the results. About 20% of the PPG signal from the nightly recordings was further eliminated due to loss of signal, primarily from the finger sensor. However, the high number of vasoconstriction events detected within each night of the available data adequately powers the correlation analyses and improves the robustness of the data. Both devices were not synchronized on the same time scale; therefore, we attempted to perform cross-correlation to 15-minute segments within a +30 or –30 second window. This technique can result in segment selection errors and lower correlation values. Future studies with concurrent use of the wristband device during clinical sleep studies with finger PPG sensors can be used to validate these results in a larger population. While we have shared preliminary results on the relationship between Mvasoc and daily pain intensity, the study is primarily focused on the ability to detect vasoconstriction events with wearable devices, and we were unable to derive any clinical inference on any causal relationship with VOC onset, the impact of hemoglobin levels, or the use of hydroxyurea on vasoreactivity.

Funding

This work was supported by grants from the National Institutes of Health; the National Heart, Lung, and Blood Institute (grant U01 HL117718); the Ming Hsieh Institute for Research of Engineering-Medicine for Cancer; and the Saban Research Institute at Children’s Hospital Los Angeles.

Authors' Contributions

PS collected and analyzed the data and wrote the manuscript. SS and MC analyzed the data. MCKK and TDC revised the manuscript. SV designed the study, collected and analyzed data, and revised the manuscript.

Conflicts of Interest

None declared.

References

1. Kato GJ, Piel FB, Reid CD, et al. Sickle cell disease. Nat Rev Dis Primers 2018 Mar 15;4:18010. [doi: 10.1038/nrdp.2018.10] [Medline: 29542687]
2. Coates TD, Chalacheva P, Zeltzer L, Khoo MC. Autonomic nervous system involvement in sickle cell disease. Clin Hemorheol Microcirc 2018;68(2-3):251-262. [doi: 10.3233/CH-189011] [Medline: 29614636]
3. Shah P, Khaleel M, Thuptimdang W, et al. Mental stress causes vasoconstriction in subjects with sickle cell disease and in normal controls. Haematologica 2020 Jan;105(1):83-90. [doi: 10.3324/haematol.2018.211391] [Medline: 30975906]
4. Veluswamy S, Shah P, Khaleel M, et al. Progressive vasoconstriction with sequential thermal stimulation indicates vascular dysautonomia in sickle cell disease. Blood 2020 Sep 3;136(10):1191-1200. [doi: 10.1182/blood.2020005045] [Medline: 32518948]
5. Khaleel M, Puliyel M, Shah P, et al. Individuals with sickle cell disease have a significantly greater vasoconstriction response to thermal pain than controls and have significant vasoconstriction in response to anticipation of pain. Am J Hematol 2017 Nov;92(11):1137-1145. [doi: 10.1002/ajh.24858] [Medline: 28707371]
6. Sangkatumvong S, Khoo MC, Kato R, et al. Peripheral vasoconstriction and abnormal parasympathetic response to sighs and transient hypoxia in sickle cell disease. Am J Respir Crit Care Med 2011 Aug 15;184(4):474-481. [doi: 10.1164/rccm.201103-0537OC] [Medline: 21616995]
7. Sangkatumvong S, Khoo MC, Coates TD. Abnormal cardiac autonomic control in sickle cell disease following transient hypoxia. Annu Int Conf IEEE Eng Med Biol Soc 2008;2008:1996-1999. [doi: 10.1109/IEMBS.2008.4649581] [Medline: 19163084]

8. Chalacheva P, Ji Y, Rosen CL, DeBaun MR, Khoo MC, Coates TD. Nocturnal peripheral vasoconstriction predicts the frequency of severe acute pain episodes in children with sickle cell disease. Am J Hematol 2021 Jan;96(1):60-68. [doi: 10.1002/ajh.26014] [Medline: 33027545]
9. Lubitz SA, Faranesh AZ, Atlas SJ, et al. Rationale and design of a large population study to validate software for the assessment of atrial fibrillation from data acquired by a consumer tracker or smartwatch: the Fitbit heart study. Am Heart J 2021 Aug;238:16-26. [doi: 10.1016/j.ahj.2021.04.003] [Medline: 33865810]
10. Kinnunen H, Rantanen A, Kenttä T, Koskimäki H. Feasible assessment of recovery and cardiovascular health: accuracy of nocturnal HR and HRV assessed via ring PPG in comparison to medical grade ECG. Physiol Meas 2020 May 7;41(4):04NT01. [doi: 10.1088/1361-6579/ab840a] [Medline: 32217820]
11. de Zambotti M, Rosas L, Colrain IM, Baker FC. The sleep of the ring: comparison of the ŌURA sleep tracker against polysomnography. Behav Sleep Med 2019;17(2):124-136. [doi: 10.1080/15402002.2017.1300587] [Medline: 28323455]
12. Mekhael M, Ho C, Noujaim C, et al. Compliance challenges in a longitudinal COVID-19 cohort using wearables for continuous monitoring: observational study. J Med Internet Res 2023 Apr 5;25:e43134. [doi: 10.2196/43134] [Medline: 36763647]
13. Biostrap. URL: https://biostrap.com/ [accessed 2024-03-21]
14. Alice NightOne. Philips. URL: https://www.documents.philips.com/doclib/enc/17329407/Alice_NightOne_Data_Sheet.pdf [accessed 2024-03-21]
15. Kemp B, Värri A, Rosa AC, Nielsen KD, Gade J. A simple format for exchange of digitized polygraphic recordings. Electroencephalogr Clin Neurophysiol 1992 May;82(5):391-393. [doi: 10.1016/0013-4694(92)90009-7]
16. Derrick TR, Thomas JM. Time-series analysis: the cross-correlation function. In: Stergiou N, editor. Innovative Analyses of Human Movement: Human Kinetics Publishers; 2004:189-205.
17. Characteristics of inpatient hospital stays involving sickle cell disease, 2000-2016. Agency for Healthcare Research and Quality. 2019. URL: https://hcup-us.ahrq.gov/reports/statbriefs/sb251-Sickle-Cell-Disease-Stays-2016.jsp [accessed 2024-03-21]
18. Espinoza J, Shah P, Veluswamy S, et al. Aura and mental stress are associated with reports of pain in sickle cell disease-a pilot study using a mobile application. Am J Hematol 2020 Apr;95(4):E101-E103. [doi: 10.1002/ajh.25736] [Medline: 31957078]
19. Veluswamy S, Shah P, Sy S, et al. Nocturnal autonomic and vasoconstriction parameters by wearable wrist photoplethysmography predict intensity of imminent pain in subjects with sickle cell disease. Blood 2023 Nov 2;142(Supplement 1):3875. [doi: 10.1182/blood-2023-187353]

Abbreviations

ANO: AliceNightOne
ANS: autonomic nervous system
Mvasoc: magnitude of vasoconstriction
PPG: photoplethysmography
PPGamp: PPG amplitude
PPGampCV: PPGamp coefficient of variation
SCD: sickle cell disease
SpO2: oxygen saturation
VOC: vaso-occlusive crises

Edited by A Malik; submitted 03.Apr.2025; peer-reviewed by J Azai, NI Shuzan; revised version received 15.Aug.2025; accepted 01.Dec.2025; published 06.Jan.2026.

Please cite as:
Shah P, Sy S, Chen M, Khoo MCK, Coates TD, Veluswamy S
Wearable Device Photoplethysmography As a Viable Tool to Longitudinally Monitor Vasoconstriction Biomarkers for Predicting Vaso-Occlusive Crisis in Sickle Cell Disease: Feasibility and Validation Study
JMIR Hum Factors 2026;13:e75465
URL: https://humanfactors.jmir.org/2026/1/e75465
doi:10.2196/75465

Original Paper

Opportunities for Improved Device Design Based on Central Line Placement Practices: Contextual Inquiry Study

Mary Beth Privitera[1], BSID, MDES, PhD; Sameer Khan[2], MD; Bilal Irfan[3,4], MS; Shayan Ali[5,6]; Cecelia Arredondo[7], MDES; Kyrsten Sanderson[7], MDES; Jordan Bonomo[8], MD

[1]University of Cincinnati, Know Why Design, LLC, Mason, OH, United States
[2]Department of Critical Care Medicine, Hoag Hospital, Newport Beach, CA, United States
[3]Department of Epidemiology, Department of Neurology, University of Michigan Medical School, Ann Arbor, MI, United States
[4]Center for Surgery and Public Health, Brigham & Women’s Hospital, Center for Bioethics, Harvard Medical School, Boston, MA, United States
[5]Neuroscience Department, Austin College, Sherman, TX, United States
[6]North Texas Medical Research Institute, Rockwell, TX, United States
[7]School of Design, College of Design, Art, Architecture & Planning, University of Cincinnati, Cincinnati, OH, United States
[8]Section Chief, Critical Care, Department of Emergency Medicine, Neurology, NeuroCritical Care, University of Cincinnati, Cincinnati, OH, United States

Corresponding Author:
Mary Beth Privitera, BSID, MDES, PhD
University of Cincinnati
Know Why Design, LLC
4691 Saddletop Ridge Lane
Mason, OH, 45040
United States
Phone: 1 5132585824
Email: privitmb@ucmail.uc.edu

Abstract

Background: Central venous catheters (CVCs) are indispensable to contemporary critical care, perioperative management, and emergency resuscitation, yet their insertion remains fraught with preventable harm and inefficiency.

Objective: This study aimed to identify all areas of CVC placement that can be improved through device design using human-centered design and qualitative research methods.

Methods: This qualitative study was a contextual inquiry of CVC placement, which included observation alongside brief face-to-face interviews with physicians. It was aimed at providing a depth of understanding using evidence to demonstrate causality. This study was conducted at 3 hospitals in the emergency department, the intensive care unit, and the operating rooms. Where possible and with additional consent, sessions were recorded in video or still photography, or at times both. This study included 19 observations and 24 interviews.

Results: In this study, the approach to CVC insertion was consistent across hospitals and care environments, with moderate variability spanning a few sections, such as suture and dressing use or lack thereof in specific care environments. The described and observed difficulties leave room for improvement in device design. The results of this study indicated that there are 34 discrete steps to placing a CVC line, with most time spent during sterile preparation. As a result of the device or kit design, challenges were observed. These included missing essential materials from kits, difficulty distinguishing between nonsterile and sterile items, challenges with lidocaine ampules, patient claustrophobia from draping, and a lack of user preference for kit contents. Additional challenges included obscured ultrasound views, kinked guidewires, overall procedural untidiness, and considerable waste management issues.

Conclusions: An intuitive kit that aligns with predictable human behavior and eliminates unnecessary multistep detours can reduce novice failure rates, cognitive load, and practice inconsistency, and it could also curb nonrecyclable waste from “backup” kits opened for a single missing item. By reframing CVC systems as sociotechnical solutions rather than static assortments of parts, the same design moves that minimize improvisation and coordination errors for physicians may also reduce dwell time and manipulation events for patients, thereby advancing the core triad of safety, procedural efficacy, and everyday usability. By examining how clinicians place central lines, this study reveals modifiable design flaws that perpetuate risk despite decades of procedural standardization. Contextual inquiry provides the evidentiary bridge between clinical imperatives to reduce complications and the practical realities of device use. Embedding such investigations at the outset of design and iteratively throughout product life cycles offers a path toward safer, more efficient, and more humane central venous access for both patients and providers.

(JMIR Hum Factors 2026;13:e84621) doi:10.2196/84621

KEYWORDS

contextual inquiry; central line placement; medical device design; patient safety; ease of use

Introduction

Central venous catheters (CVCs) are indispensable to contemporary critical care, perioperative management, and emergency resuscitation, yet their insertion remains fraught with preventable harm and inefficiency. Large contemporary syntheses estimate that approximately 30 in every 1000 patients with a CVC in place for 3 days will experience at least one serious complication, be it arterial cannulation, pneumothorax, central line–associated bloodstream infection (CLABSI), or deep venous thrombosis, whereas catheter malfunction alone occurs at roughly 6 events per 1000 catheter days [1]. Additionally, the insertion process can be a source of complication. CLABSIs still claim thousands of lives each year in the United States and add billions of dollars to health care costs despite decades of prevention bundles and increasingly routine ultrasound guidance intended to curb mechanical mishaps and accelerate successful cannulation [2].

While epidemiological surveillance and randomized trials may have helped refine insertion checklists and sterility protocols, far less attention has been paid to the concrete interactions among users, tools, and environments that shape everyday practice. Human factors frameworks may be an impetus to study actual work as done rather than work as imagined. The US Food and Drug Administration’s 2016 guidance on applying human factors in medical devices and the Association for the Advancement of Medical Instrumentation TIR51 standard on contextual inquiry may provide the grounding for early, field-based observation and interview techniques to surface latent use errors, cognitive burdens, and design mismatches before devices reach market or are iterated for safety [3,4].

This study applied a structured contextual inquiry across 3 hospitals and multiple care settings to showcase how central line placement unfolds in real-world settings minute by minute, from presterile preparation to dressing application. By triangulating in situ observation with brief clinician interviews, 34 discrete steps were catalogued, mapping procedure time distributions and documenting recurrent friction points that compromise sterility, ergonomics, and situational awareness. This study aimed to identify all areas of CVC placement that can be improved through device design. From the drapes that obscure patients’ faces and guidewires that become kinked to the missing of essential or preferred components, this study highlights the basis for next-generation kit architecture, accessory design, and room layout that align with real user needs in both chaotic and controlled environments.

Methods

Overview

This qualitative study was a contextual inquiry of CVC placement, which included observation alongside brief face-to-face interviews with physicians. It was aimed at providing a depth of understanding using evidence to demonstrate causality. According to Maxwell [5,6], qualitative research is well suited for causal inference as it allows for detailed examination of specific processes and mechanisms in real-world contexts, revealing how and why outcomes occur beyond mere correlations [5,7]. A detailed description of the process, tools, and people involved is provided below. The technique of contextual inquiry is promoted by the Food and Drug Administration human factors guidance (2016) to determine user needs at the start of any design process. The data collected in this study included both observational and interview data, with 3 main areas of focus: the user, the environment, and the tasks as part of the steps and workflow.

This study was conducted at 3 hospitals in the emergency department (ED), intensive care unit (ICU), and operating rooms (ORs). Where possible and with additional consent, sessions were recorded on video or still photography, or at times both. This study included 19 observations and 24 interviews (Table 1). The total number of observations and interviews conducted at each site was determined by the clinical need and patient-clinician consent at the time of the study. In instances in which patient consent was not provided but physicians wanted to participate, the physicians could opt to demonstrate their placement technique using a simulated patient in the care environment.

Table 1. Breakdown of observations and interviews by location.

	Cincinnati, Ohio, n (%)	Wake Forest, North Carolina, n (%)	Sacramento, California, n (%)
Observations (n=19)	9 (47)	7 (37)	3 (16)
Interviews (n=24)	8 (33)	8 (33)	8 (33)

A total of 24 physicians were involved in this study, including emergency medicine, critical care, and anesthesiology specialists. Observations and interviews were conducted in perioperative rooms, ORs, EDs, and ICUs.

Within each hospital, the same brand of kit was used; however, in one instance at Wake Forest, there was a different type of kit used between the OR and the ICU and ED. The kits consisted of other brands and product identification numbers (Table 2).

Table 2. Brand of kits used by location and environment of care.

	Cincinnati, Ohio	Wake Forest, North Carolina	Sacramento, California
Operating room	Edwards Multi-Med CVC[a] 3K20N18141NL	Arrow central venous access kit ASK-21242-PCMH1	Arrow Blue Plus pressure-injectable multilumen CVC kit ASK-45703-PIO
Intensive care unit	Edwards Multi-Med CVC 3K20N18141NL	Arrow pressure-injectable multilumen CVC kit CDC-45703-XP1A	Arrow Blue Plus pressure-injectable multilumen CVC kit ASK-45703-PIO
Emergency department	Edwards Multi-Med CVC 3K20N18141NL	Arrow pressure-injectable multilumen CVC kit CDC-45703-XP1A	Arrow Blue Plus pressure-injectable multilumen CVC kit ASK-45703-PIO

[a] CVC: central venous catheter.

All interviews followed the core principles of contextual inquiry studies in that they were conducted in the users' real-world environment, with the research team establishing a master (physician)-apprentice (researcher) relationship. As the procedure unfolded, and at opportune times, the researcher would share interpretations to uncover observations and deeper insights. All the interviews were guided by the tasks in the CVC placement procedure and focused on the usability of each element. Typical observation or interview experiences ranged from 45 minutes to 1.5 hours, as influenced by clinical responsibilities and the participants' availability.

Ethical Considerations

The study protocol was appropriately reviewed by the University of Cincinnati institutional review board, which determined that it did not constitute human participant research under federal regulations, as the primary focus was on device improvements and quality of care rather than generating generalizable knowledge about individuals. Despite the nonresearch classification, informed consent was required from all participants. As detailed observations were conducted of clinical procedures that could capture sensitive patient information or professional performance details, robust measures were in place to protect privacy and confidentiality. In addition, efforts were made to ensure a diverse participant pool to avoid biased insights that could perpetuate disparities in device design or procedural improvements.

Results

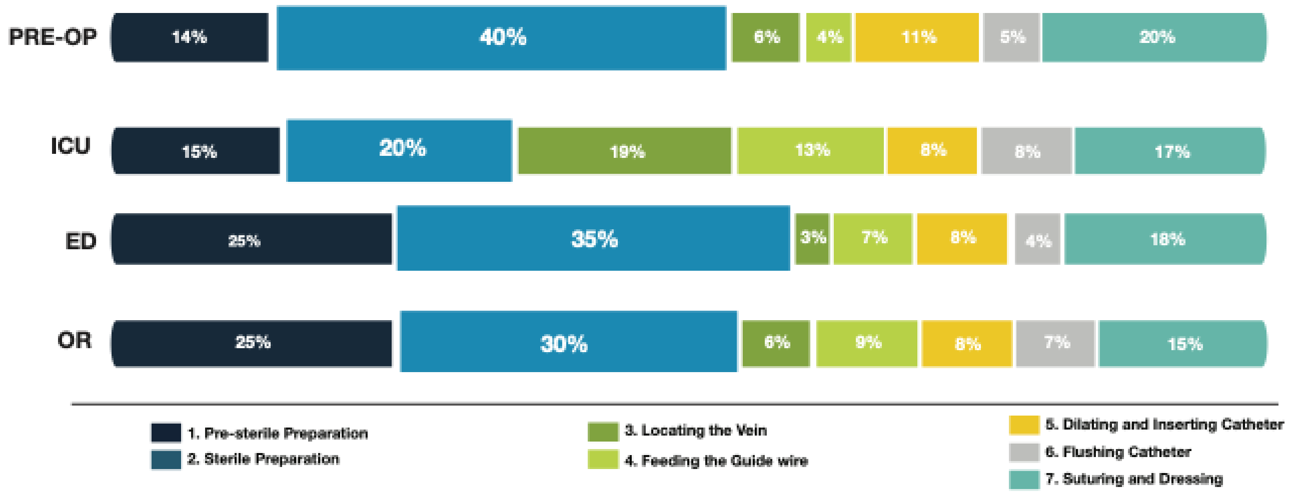
In this study, the average overall time to place a CVC varied across care environments (Table 3). In discussions with clinicians, some postulated that procedure times vary proportionally with the provider's perceived control over the patient and environment (ie, that the more chaotic the environment, the longer the procedure).

Table 3. Average central venous catheter (CVC) placement timing in different environments of care.

Environment	CVC placement time
Operating room	16 min, 48 s
Preoperative care	20 min, 15 s
Intensive care unit	27 min, 53 s
Emergency department	35 min, 28 s
Overall, mean (SD)	25 min, 6 s (7 min 12 s)

Procedural timing was also broken down into 7 distinct sections within the overall procedure (Figure 1). These included the following: presterile preparation, sterile preparation, vein localization, guidewire insertion, catheter dilation and insertion, catheter flushing, and suturing and dressing. Of note, time from presterile to sterile preparation accounted for more than half of the overall procedure time in all locations except the ICU. The research team noted that the number of staff members participating in setting up for the procedure varied across sites and care environments.

Figure 1. Percentage of time per section of the central venous catheter placement procedure by environment. ED: emergency department; ICU: intensive care unit; OR: operating room.



The patient experience during CVC placement highlights conditions that impact overall safety, efficacy, and usability. Draping a conscious patient can be problematic. During procedure preparation, the provider unpacks sterile drapes from a kit and prepares a sterile field directly over the patient (Figure 2). This extends the sterile field, albeit an unstable one, directly over the patient. As a result of patient movement, the required equipment for the procedure can fall out of the sterile field or become lost. Uncooperative patients require additional procedures, such as intubation, muscle relaxation, and sedation, before CVC placement.

Figure 2. Patient acting as an extended sterile work surface.

In the OR, the patients were intubated and unconscious. However, in the ED and ICU environments, the patients were often conscious and mobile. In all cases, draping covered the patient's face throughout the procedure, creating a potentially claustrophobic and challenging environment (Figure 3). In clinical practice, there is substantial variability in body habitus across the patient population. As a result, the drapes provided in the kits may not fit patients classified as obese. In some cases, the aperture provided by the drape is not large enough to accommodate the insertion location, resulting in the provider improvising by creatively engineering a larger opening.

Figure 3. A patient drape placed over the patient's face in the intensive care unit.

Patient airways require management throughout the procedure. In this study, patients in the ICU were fitted with an oxygen mask. A mask is often placed in case a future procedure is needed to avoid disturbing the sterile field.

Gaining access and positioning can often lead to patient discomfort. Patients are required to hold awkward Trendelenburg positions with their head turned and extended to optimize their anatomy for successful CVC placement (Figure 4). Patients were instructed to maintain position throughout the procedure, although it can be uncomfortable because the plastic portion of the drape may be directly on a patient's skin. For conscious patients, this at times required further anesthesia and posed challenges for physicians, given that, as access was gained, patients would move. This position proves challenging for patients who are obese who may also require additional supplies and more extended periods to complete procedure preparation.

Figure 4. Aperture opening of the drape adhered directly to a patient's face.

The provider experience during CVC placement also underlines conditions that could impact overall safety, efficacy, and usability. Physicians are required to don personal protective equipment in tight spaces without compromising sterility (Figure 5). In this study, the order of preparation varied across care locations. For example, ED physicians sometimes elected to preclean the patient before gowning, which allowed time for the chlorhexidine preparation to dry. In contrast, in the OR, some physicians prepared personal protective equipment in advance.

Figure 5. Donning of personal protective equipment in a tight space.

In the CVC procedure, 34 individual steps require the provider to locate key pieces of equipment and determine where and how to use them while maintaining a sterile workspace. During this multistep process, physicians often improvise, which may place a heavy cognitive burden on them and require coordination. For example, physicians will think through the procedure before beginning, questioning whether they have all the necessary supplies and where to place them to maximize access while maintaining the sterile field and coordinating with other support personnel.

Across all sites and use environments, at least 2 physicians were present. This enabled 1 sterile operator and 1 support person to adjust patient drapes, including using a towel to protect the patient's face; retrieve items that were required but not readily available in the kit or forgotten; and sterilely prepare the ultrasound probe.

Physicians were often in awkward positions due to suboptimal biomechanics resulting from product design, room layout, or both. For example, physicians struggled to fit all the necessary equipment on a single surface due to space constraints. The equipment is often split between multiple work surfaces (ie, patient and table), forcing the provider to turn to access both surfaces and remove their focus from the patient. The space constraints also make it difficult to avoid touching a nonsterile surface during the procedure. Furthermore, flushing with a standard syringe requires an uncomfortable and awkward hand movement needed to achieve aspiration, as the device is designed to deliver a solution. When ultrasound is used for safe access, both hands are required to hold the equipment, which can make aspiration a challenging one-handed maneuver (Figure 6).

Figure 6. One-handed aspiration while holding the ultrasound probe.

Physicians held the syringe and introducer needle at a 45 to 30° angle and subsequently lowered the angle while advancing the needle (Figure 7). To verify position, physicians would draw more blood and assess color and pressure gradient. Some commented that some step was the highest "stress" point in the procedure. One intensivist even remarked that finding the vein was the most stressful part and that, once the wire was in and the needle was out, their stress subsided. At this point, there is a risk of pneumothorax or perforation of the carotid artery. It was observed that maintaining stability and aspirating to visualize venous blood return is ergonomically challenging for physicians. This is further complicated by the ultrasound monitor being placed outside the provider's field of view (Figure 8).

Figure 7. Access needle angle relative to the ultrasound probe.

Figure 8. Ultrasound monitor position requires physicians to look away from the access site.

Additionally, the pressure that the user applies to the vein with the ultrasound probe can alter central venous anatomy when the device is withdrawn from the skin surface. Physicians were at times compelled to make slow, careful movements, with risk controlled by the user's experience. Rather than an intuitive device design, advanced troubleshooting or error prevention during the procedure may depend on the provider's expertise and dexterity.

While simultaneously troubleshooting and preventing errors, physicians are exposed to sharps and biohazards once access to the vein is achieved. For example, users are exposed to the access needle, where the skin incision scalpel is typically contaminated with the patient's blood within the working area (Figure 9).

Figure 9. Opening the kit midprocedure with the sharps anchored in the kit.

While efforts were made to control and maintain the safety of sharps, the CVC kit design places the cognitive burden, safety, and responsibility on the provider. As vein dilation occurs, the access site often becomes highly disordered and poses a biohazard to health care physicians (Figure 10). Once the vein was dilated, increased blood flow increased the stress and inconvenience of the procedure in that the guidewire itself became a slippery surface. It could become misplaced, requiring additional intervention. This hazardous situation also extends to catheter insertion (Figure 11).

Physicians were required to thread the catheter over the guidewire and adopted a coiling behavior to maintain control and reduce proximal tip movement, which can complicate device alignment (Figure 12).

Figure 10. Advancement of the dilator over the wire in slippery conditions due to venous return.

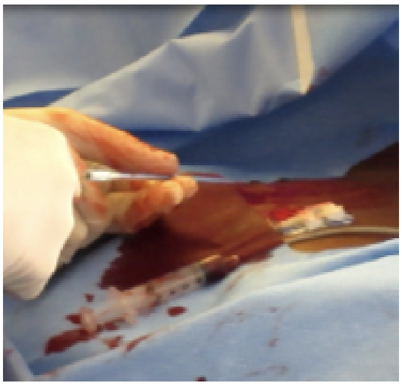


Figure 11. Catheter insertion with exposure to biohazard.

Figure 12. Coiled guidewire enabling control during catheter insertion.

These situations were particularly challenging for novice health care providers, who grappled with each step of the procedure (Figure 13). Consequently, inexperienced health care physicians often have unsuccessful attempts to place CVCs.

The product design of traditional CVC kits was not always entirely optimized for streamlined assembly. Many kits, for example, did not have all the supplies or the preferred supplies for CVC placement. As a result, clinicians often spent additional time gathering the necessary materials (eg, skin preparation, gauze, saline, and lidocaine) as well as the preferred materials because the valves provided in the kits were at times not compatible with other hospital equipment, thus suggesting better packaging so that the risk or delay in the procedure is minimized. Some users preferred vertical valves such as clave connectors (Figure 14). Physicians who chose to use nonstandard or additional supplies often opened a secondary kit sourced from separate locations within the hospital. Items collected may include additional drapes, Mayo stand, lidocaine, sterile flush syringes, chlorhexidine preparation, and ultrasound devices or supplies.

When physicians remove the guidewire, there is no proper receptacle to place it in; it is thrown away with nonsharp materials or hazardous materials. In this study, some physicians did not use sharps holders during the procedure, preferring to leave sharps on the tray (Figure 15), suggesting an unmet need among users. For those who used the temporary sharps holder provided in some trays, it was poorly secured and became top-heavy, often leading to dropped sharps. Additionally, there were no safeguards for the disposal of all sharps.

Figure 13. Novice users have increased exposure to biohazards.

Figure 14. Different forms of valve connectors.

Figure 15. Sharps are placed on top of the kit during the procedure.

There is a considerable amount of nonrecyclable waste and challenges because of the kit design (Figure 16). In this study, the kits at times lacked a clear delineation between the layers that designate items and specify when they should be used. There were excess materials not used in the procedure, which contributed to the overall waste. Physicians often draped the sterile area and manipulated the drape around the patient's face and head, making it easier for the patient to breathe and reducing claustrophobia.

Figure 16. Waste accumulation because of the kit design.

The quality and design of the CVC products' individual elements pose usability challenges, including misalignment with predictable human behavior, heavy reliance on memory, and a lack of visual or tactile references. One challenge in this study was that the guidewire kinked frequently. When this happened, physicians would often try to remove the kink or use the wire regardless. If this guidewire is rendered unusable, the provider may open a new kit to replace the failed one or, in some instances, a missing part (Figure 17).

Figure 17. Three kits opened for the guidewire and dilator due to kinking.

In this study, the dilator was the same color as the draping and could be camouflaged, eventually getting lost in it, causing health care providers to search. Additionally, due to the complex setup, components sometimes fell outside of the sterile field. Another detail that caused some difficulty in the guidewire design was the "J" curve. While it can provide safe intravascular advancement, this curve on the wire can also make the guidewire more difficult to thread into the catheter. Some physicians were observed making what some may describe as counterintuitive workarounds and flipping the guidewire to the straight end to easily advance it through the catheter. Though well intended, this variation in CVC placement may increase the risk of complications.

Furthermore, there is little to no tactile feedback from the guidewire feeder, making it difficult for the user to know whether the wire is advancing. This is further complicated by blood-stained gloves, which make the guidewire markings difficult to see and grasp. The results of these outdated design decisions can create an unnecessary burden on physicians to judge catheter alignment on the guidewire and insertion depth.

Discussion

Principal Findings

In this study, the approach to CVC insertion was consistent across hospitals and care environments, with moderate variability spanning a few sections, such as suture and dressing use or lack thereof in specific care environments. The described and observed difficulties leave room for improvement in device design (Table 4).

Table 4. Summary of use-related issues throughout the procedure.

Procedure section and device used	Problem area
Presterile preparation	
Central venous catheter kit	• It does not include all essential materials needed for central line placement, resulting in unnecessary delays. • It is unable to distinguish nonsterile and sterile items during nonsterile preparation.
Nonsterile syringe	• Health care providers typically empty 3-5 syringes (30-50 mL) of saline solution into the kit basin from nonsterile saline syringes. These were not included in the kit.
Lidocaine	• Because the lidocaine is not sterile, health care providers will either have someone hold it for them, tape it to a Mayo cart, or chase down the ampule while loading the syringe.
Sterile preparation	
Patient drape	• Due to the placement of the full-body drape, patient claustrophobia can occur; it is challenging for the provider to maintain sterility and manage the patient during the procedure.
Locating the vein	
Introducer needle	• Physicians expressed a preference for the 18-gauge introducer needle without a catheter over the 20-gauge needle with an 18-gauge catheter assembly. For patients with obesity, physicians noted that the length (2.5 inches) of the needle was not sufficient. The introducer needles were not echogenic; thus, they were difficult to view in the ultrasound monitor.
Ultrasound	• The ultrasound monitor was not placed within the line of sight of the provider, making it difficult for them to effectively use the tool. • Health care providers do not generally document ultrasound use by capturing an image of the procedure as required by major insurers for reimbursement.
Feeding the guidewire	
Guidewire	• Wire kinking while executing the procedure was problematic, and once deemed unusable, the provider would open another kit only to access an additional wire. • The “J” tip of the wire was difficult to thread and feed. • Physicians had difficulty estimating how much of the wire was inserted. If they were unsure, they took the wire out of the delivery system. Once the wire is removed from the delivery system, it is difficult to reload; thus, once it is out, it typically stays out. Markings on the wire were also not clear to the providers. • Due to the inherent messiness of the procedure, blood is often on the providers’ hands, making it difficult to handle the guidewire delivery system.
Dilating and inserting the catheter	
Dilator	• The dilator often kinks if the provider aggressively pushes. If the dilator is no longer usable, another kit is opened only to access another dilator. • The dilator length (4 inches) is not sufficient for patients with obesity.
Catheter	• Threading the catheter over the wire is a difficult task. Physicians have a hand tremor when performing this task. • It is difficult for the provider to determine whether the tip of the catheter has reached the superior vena cava. Physicians will estimate placement and follow up by x-raying the patient to confirm proper placement.
Flushing the catheter	
Syringe	• Physicians faced challenges with maintaining proper syringe grip during aspiration and flushing. Physicians struggle with the current syringe design to alternate between 2 grips.
Valves	• The T-shaped valves provided in the kit also present challenges when flushing the line after placement. The syringe often slips off the valve when flushing the lines.
Suturing and dressing placement	
Suture loop (feature) on the catheter	• The suture loop anchor and box clamp were not always used. If there is excess catheter outside of the patient’s body, depending on the insertion location and the patient’s anatomy, the provider will use the suture loop anchor and box clamp.

This contextual inquiry reveals that central venous catheterization is not merely a sequence of technically codified steps but also an intricate sociotechnical performance in which device design, environmental constraints, and clinician improvisation intersect to shape risk. Our mapping of 34 discrete actions, the dominance of preparation time in most settings, and the recurrent need to compensate for missing or poorly engineered components align with contemporary estimates that roughly 3% of patients exposed to a catheter for 3 days sustain a major complication [1]. Large contemporary reviews showcase that CVC is now routine, roughly 8% of hospitalized patients need a CVC, and more than 5 million are placed annually in the United States. Ultrasound-guided puncture should be the default. However, Kehagias et al [8] found that observations of obscured monitors, awkward one-handed aspiration, and improvised wire handling show how the safest technique on paper can still be undermined by poor ergonomics and kit design.

Our findings build on imperatives for field-based inquiry into real work conditions to surface latent design hazards. We observed exactly the kinds of use errors that are of concern: kinking guidewires without tactile feedback, dilators camouflaged against drapes, valves incompatible with existing hospital hardware, and ultrasound displays positioned outside the operator's natural sightline. Such mismatches between device affordances and predictable human behavior may shift cognitive burden to clinicians, who must remember workarounds and coordinate ad hoc assistance while maintaining sterility. A contextual inquiry approach is designed to elicit these mismatches; our study demonstrates its value in an acute, invasive procedure where seconds and millimeters matter. The imperative to integrate these insights into formal design controls is clear.

Several concrete implications for device and kit redesign emerge from our contextual findings and accord with human factors guidance that interventions must fit users' capabilities, workflows, and environments to sustain adoption and fidelity [9-12]. The suggested improvements and justifications are mentioned subsequently.

First, rather than perpetuating one-size-fits-all assortments, kits could be organized and sufficiently complete for the intended procedure and patient mix, thus minimizing secondary searches and redundant openings so that routinely needed sterile flushing media, local anesthetic, echogenic introducer needles of adequate length for patients classified as obese, and valve connectors compatible with local infusion hardware become immediately available; these specific component choices derive from our observations, whereas the general mandate to reduce cognitive and physical load is articulated in human-centered design literature [9,12].

Second, because the procedure advances in a sequential order, CVC kit components should be packaged in the same order to reduce cognitive burden. High-risk elements could incorporate salient visual and tactile affordances such as blood-tolerant depth markings on guidewires, textured feeders that signal advancement, and color-contrasted dilators that cannot visually disappear against drapes, an approach consistent with human factors recommendations to engineer cues that support rapid, accurate action under stress [12].

Third, ultrasound ergonomics require particular attention: practice reviews emphasize that the operator should keep the puncture site, needle, and image within a single line of sight, yet we repeatedly observed monitors positioned laterally or behind the user, encouraging awkward posture and one-handed aspiration [9-11].

Fourth, because maintaining stability of the introducer needle at the insertion site is critical, there is an opportunity for a device that minimizes hand movement or stabilizes the ultrasound probe during vessel access.

Fifth, converging evidence from simulation and ergonomic assessments shows that suboptimal screen and table positioning increases musculoskeletal strain and facilitates needle advancement errors, especially among novices, supporting the design of articulating mounts, probe-holding accessories, and workstation layouts that free the dominant hand and keep the image within the operative field [10-12].

Sixth, additionally, as insertion components are introduced sequentially, combining the components associated with adjacent actions into a single integrated unit would significantly simplify the procedure.

Seventh, as threading the dilator over the guidewire is related to a risk of kinking, a device may reduce or eliminate the need for over-the-wire threading while reducing the risk of guidewire kinking.

Eighth, patient experience is also inseparable from considerations of safety. Drapes that blanket a conscious patient's face created anxiety and impeded airway access, whereas Trendelenburg positioning with head rotation was difficult to sustain, particularly for individuals classified as obese, who then required additional supplies and time. Our observations suggest that rethinking drape architecture to permit facial exposure without compromising sterility and integrating oxygen delivery ports or transparent windows could mitigate claustrophobia and facilitate airway monitoring.

Because CLABSI remains among the costliest hospital-acquired infections on a per-case basis, the marginal gains from such design tweaks may translate into substantial economic and human benefit when scaled across millions of annual insertions [13]. Even in the era of prevention bundles, catheter dwell time remains a potent infection driver; CLABSI rates in one 2-year adult cohort climbed from 4.80 to 8.64 per 1000 catheter days as dwell exceeded 20 days, with multidrug resistant *Acinetobacter baumannii* predominating, mirroring some of our concerns that missing components, repeated kit openings, and ad hoc maintenance steps prolong setup and line life, thereby compounding exposure to contamination [14]. It has been previously emphasized that complication profiles can hinge on site choice, catheter caliber, and positioning—adult data (eg, 3SITES study) link subclavian access to fewer infections but more pneumothoraxes, whereas pediatric series show different risk patterns—and it has been reiterated that neutral rather than exaggerated positioning and meticulous ultrasound use can reduce failed passes and arterial hits, resonating with our field notes on stressful needle advancement angles, off-axis screens, and repeated punctures at the point deemed to be of the highest stress [15]. Other recommendations also somewhat align with our design proposals: select the smallest necessary lumen count;

favor nontunneled catheters for fewer than 3 to 4 weeks and peripherally inserted central catheters when therapy exceeds 6 days; cap prolonged catheterization (approximately 14 days) to curb bacteremia; and mandate real-time ultrasound, chlorhexidine alcohol preparation, and daily site surveillance [16]. These were practices that our contextual inquiry somewhat related to and found were variably executed or actively hindered by kit incompleteness, unclear layer sequencing, and drapes that compromise both sterility and patient comfort. In a simulation of 40 anesthesia providers, patient safety (mean importance score 83.9/100), ease of use (mean score 64.6/100), and reduced clinician error (mean score 61.1/100) topped the selection criteria. A novel CVC system with a sequentially organized tray, enhanced labeling, and a guidewire funnel earned significantly higher scores for satisfaction overall, ease of use, layout, and safety ($P\leq.01$ in all cases) and reduced 5 of 7 common risks (including clinician error and contamination or infection), aligning with the usability deficits (wire kinking, component hunting, and ambiguous tray hierarchy) that our study catalogued [7].

This study also highlights how waste and sharps handling were downstream consequences of kit design. The absence of designated receptacles for used guidewires and scalpels, the lack of sharps holders integrated into trays, and the routine opening of multiple kits for a single missing part create biohazard exposure and substantial nonrecyclable waste. While our qualitative approach did not quantify environmental impact, other health systems work has emphasized the financial and ethical importance of reducing unnecessary disposables; therefore, future prototypes should embed closed-loop sharps to capture and minimize redundant components to support infection prevention and sustainability goals simultaneously. This inference from our data warrants targeted life cycle and cost analyses in subsequent studies.

Viewed through the intersecting lenses of safety, efficacy, and usability for patients, providers, and products, our findings provide the basis for an argument that a device ecosystem should simultaneously consider the precepts mentioned subsequently.

The first precept is to shield patients from infection, vessel injury, and the cascade of “extra” procedures (intubation and deep sedation) by shortening setup and puncture time and by making correct J-wire orientation and tip control essentially foolproof.

The second precept is to protect clinicians from sharps and blood under tension, awkward postures, and protocol drift by embedding ergonomic grips; one-handed aspiration aids; and closed, labeled receptacles that keep contaminated instruments off ad hoc trays.

The third precept is to streamline the product itself with sequential, memory-light assembly; color and texture coding; and built-in safeguards (eg, wire funnels, depth markings visible through bloodied gloves, and lockable sharps wells) that are robust and attuned for real-world conditions such as low light; urgent timelines; moving patients; and gloved, fatigued hands. In effect, an intuitive kit that aligns with predictable human behavior and eliminates unnecessary multistep detours can reduce novice failure rates, cognitive load, and practice inconsistency, just as it could serve to curb nonrecyclable waste from “backup” kits opened for a single missing item.

By reframing CVC systems as sociotechnical solutions rather than static assortments of parts, the same design moves that minimize improvisation and coordination errors for physicians also reduce dwell time and manipulation events for patients, thereby advancing the core triad of safety, procedural efficacy, and everyday usability.

Limitations

Our findings should be interpreted considering several limitations. This study was conducted in 3 US hospitals and included 19 observations and 24 interviews, which constrains generalizability, particularly to resource-limited settings or institutions that use different kit vendors. Although we captured real procedures, observer presence and video recording may have altered behavior, and some demonstrations on simulated patients cannot fully reproduce the stressors of an emergent cannulation. We did not measure clinical outcomes linked to the specific use problems we identified, nor did we quantify cognitive load or musculoskeletal strain. Furthermore, our sample comprised physicians; nurses, technicians, and infection preventionists often influence setup, maintenance, and postplacement care, and their perspectives warrant inclusion in a fuller systems analysis.

Conclusions

Future work should transition from description to intervention. Rapid-cycle prototyping informed by these contextual insights, followed by high-fidelity simulation and usability testing in accordance with relevant guidelines, could generate validated design changes. Eye tracking, motion capture, and workload assessment tools may quantify how redesigned components alter gaze patterns, posture, and error rates. Multicenter trials comparing integrated, human-centered kits with current products could measure effects on insertion time, complication rates, waste generation, and cost. Furthermore, translating these methods to maintenance phases, including line access and dressing changes, could extend human factors benefits beyond placement to the complete catheter life cycle.

By examining how clinicians place central lines, this study reveals modifiable design flaws that perpetuate risk despite decades of procedural standardization. Contextual inquiry provides the evidentiary bridge between clinical imperatives to reduce complications and the practical realities of device use. Embedding such investigations at the outset of design and iteratively throughout product life cycles offers a path toward safer, more efficient, and more humane central venous access for both patients and providers.

Acknowledgments

The authors extend their special thanks to the University of Cincinnati, Department of Emergency Medicine, for their support in making this research possible. Becton Dickinson provided funding support for this research.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy and confidentiality purposes.

Authors' Contributions

The data collection and initial analysis of this contextual inquiry research were led by JB and MBP, with senior direction from KS and CA.

Conflicts of Interest

None declared.

References

1. Teja B, Bosch NA, Diep C, Pereira TV, Mauricio P, Sklar MC, et al. Complication rates of central venous catheters: a systematic review and meta-analysis. JAMA Intern Med 2024 May 01;184(5):474-482. [doi: 10.1001/jamainternmed.2023.8232] [Medline: 38436976]
2. Central Line-associated Bloodstream Infection (CLABSI) basics. Centers for Disease Control and Prevention. 2025 Jun 12. URL: https://www.cdc.gov/clabsi/about/index.html [accessed 2025-07-24]
3. Applying human factors and usability engineering to medical devices. U.S. Food & Drug Administration. 2016 Feb. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices [accessed 2025-07-24]
4. AAMI TIR51:2014 (R2017): human factors engineering - guidance for contextual inquiry. ANSI. 2014. URL: https://webstore.ansi.org/standards/aami/aamitir512014r2017?srsltid=AfmBOopt6jPd7hRIApqzSnyRv9ZCXa_swEE4FWhr_lpW2e3e-2bm-s1B [accessed 2025-07-24]
5. Maxwell JA. Causal explanation, qualitative research, and scientific inquiry in education. Educ Res 2004 Mar 01;33(2):3-11. [doi: 10.3102/0013189x033002003]
6. Maxwell JA. A Realist Approach for Qualitative Research. Thousand Oaks, CA: SAGE Publications; 2012.
7. Wagner R, Alsbrooks K, Bacon E, O'Brien M, Hoerauf K. Clinician satisfaction and preference for central venous catheter systems promoting patient safety, ease-of-use and reduced clinician error. Expert Rev Med Devices 2023 Jul 04;20(7):607-614 [FREE Full text] [doi: 10.1080/17434440.2023.2219001] [Medline: 37272116]
8. Kehagias E, Galanakis N, Tsetis D. Central venous catheters: which, when and how. Br J Radiol 2023 Nov;96(1151):20220894 [FREE Full text] [doi: 10.1259/bjr.20220894] [Medline: 37191031]
9. Saugel B, Scheeren TW, Teboul JL. Ultrasound-guided central venous catheter placement: a structured review and recommendations for clinical practice. Crit Care 2017 Aug 28;21(1):225 [FREE Full text] [doi: 10.1186/s13054-017-1814-y] [Medline: 28844205]
10. Dzulkafli AH, Abdull Wahab SF, Othman R. Ergonomics risk assessment of musculoskeletal disorder during ultrasound-guided internal jugular venous cannulation. Malays J Med Sci 2024 Oct 08;31(5):196-204. [doi: 10.21315/mjms2024.31.5.13] [Medline: 39416728]
11. Miyazaki A, Fujii A, Kuwabara D, Minoguchi K, Kawakami H, Nakamura K, et al. Enhancing the quality and safety of central venous catheter insertion using projection mapping: a prospective observational simulation study with eye-tracking glasses. BMJ Open 2025 May 24;15(5):e095803 [FREE Full text] [doi: 10.1136/bmjopen-2024-095803] [Medline: 40413056]
12. Edwards Iii GF, Zagarese V, Tulk Jesso S, Jesso M, Harden SM, Parker SH. Designing healthcare for human use: human factors and practical considerations for the translational process. Front Health Serv 2022 Jan 6;2:981450 [FREE Full text] [doi: 10.3389/frhs.2022.981450] [Medline: 36925891]
13. Overall and unit costs of the five most common hospital-acquired infections (HAIs) in the US. One Health Trust. 2013 Nov 8. URL: https://onehealthtrust.org/publications/infographics/overall-and-unit-costs-five-most-common-hospital-acquired-infections-in-us/ [accessed 2026-01-24]
14. Pitiriga V, Bakalis J, Kampos E, Kanellopoulos P, Saroglou G, Tsakris A. Duration of central venous catheter placement and central line-associated bloodstream infections after the adoption of prevention bundles: a two-year retrospective study. Antimicrob Resist Infect Control 2022 Jul 15;11(1):96 [FREE Full text] [doi: 10.1186/s13756-022-01131-w] [Medline: 35841083]
15. Jamshidi R. Central venous catheters: indications, techniques, and complications. Semin Pediatr Surg 2019 Feb;28(1):26-32. [doi: 10.1053/j.sempedsurg.2019.01.005] [Medline: 30824130]

16. Estrada-Orozco K, Cantor-Cruz F, Larrotta-Castillo D, Díaz-Ríos S, Ruiz-Cardozo MA. Central venous catheter insertion and maintenance: evidence-based clinical recommendations. Rev Colomb Obstet Ginecol 2020 Jun 20;71(2):115-162 [FREE Full text] [doi: 10.18597/rcog.3413] [Medline: 32770871]

Abbreviations

CLABSI: central line–associated bloodstream infection
CVC: central venous catheter
ED: emergency department
ICU: intensive care unit
OR: operating room

Edited by A Kushniruk; submitted 22.Sep.2025; peer-reviewed by FNJ Frakking, I Culverhouse, R Marshall; comments to author 11.Nov.2025; revised version received 09.Jan.2026; accepted 19.Jan.2026; published 06.Feb.2026.

Please cite as:
Privitera MB, Khan S, Irfan B, Ali S, Arredondo C, Sanderson K, Bonomo J
Opportunities for Improved Device Design Based on Central Line Placement Practices: Contextual Inquiry Study
JMIR Hum Factors 2026;13:e84621
URL: https://humanfactors.jmir.org/2026/1/e84621
doi:10.2196/84621
PMID:41557444

How to Evaluate the Accuracy of Symptom Checkers and Diagnostic Decision Support Systems: Symptom Checker Accuracy Reporting Framework (SCARF)

Marvin Kopka, PhD, Dr rer medic, MPH, MSc, BSc; Markus A Feufel, Dipl-Ing (FH), MSc, PhD

Division of Ergonomics, Department of Psychology & Ergonomics (IPA), Technische Universität Berlin, Straße des 17. Juni 135, Berlin, Germany

Corresponding Author:
Marvin Kopka, PhD, Dr rer medic, MPH, MSc, BSc
Division of Ergonomics, Department of Psychology & Ergonomics (IPA), Technische Universität Berlin, Straße des 17. Juni 135, Berlin, Germany

Abstract

Symptom checkers are apps and websites that assist medical laypeople in diagnosing their symptoms and determining which course of action to take. When evaluating these tools, previous studies primarily used an approach introduced a decade ago that lacked any type of quality control. Numerous studies have criticized this approach, and several empirical studies have sought to improve specific aspects of evaluations. However, even after a decade, a high-quality methodological framework for standardizing the evaluation of symptom checkers is still lacking. This paper synthesizes empirical studies to outline the Symptom Checker Accuracy Reporting Framework (SCARF) and a corresponding checklist for standardizing evaluations based on representative case selection, an externally and internally valid evaluation design, and metrics that increase cross-study comparability. This approach is supported by several open access resources to facilitate implementation. Ultimately, it should enhance the quality and comparability of future evaluations of online and artificial intelligence (AI)–based symptom checkers, diagnostic decision support systems, and large language models to enable meta-analyses and help stakeholders make more informed decisions.

(JMIR Hum Factors 2026;13:e76168) doi:10.2196/76168

KEYWORDS

symptom checker; symptom assessment applications; evaluation; case vignettes; preclinical; decision support; large language model; data analysis; health technology assessment; impact evaluation; artificial intelligence; AI

Introduction

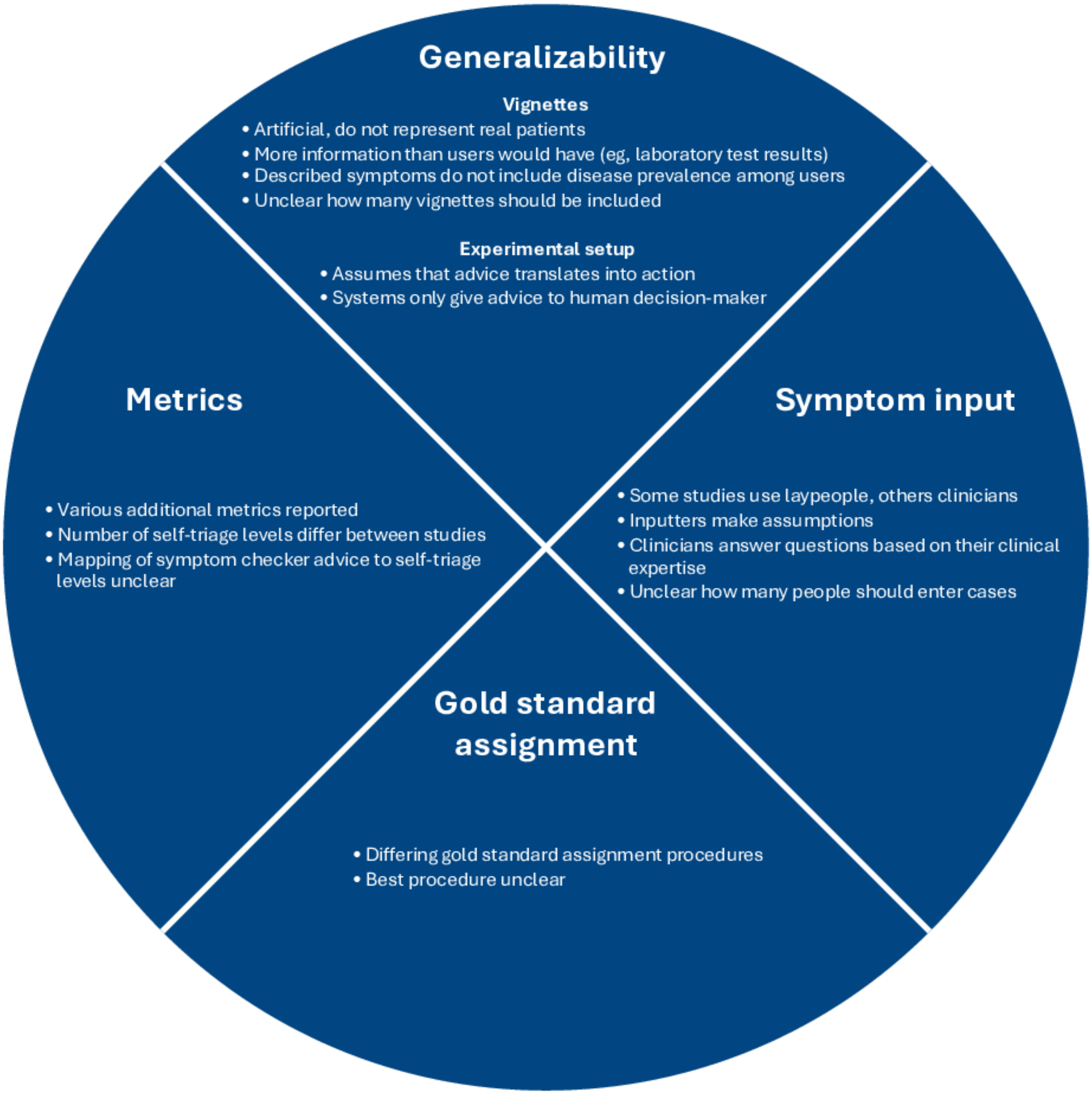
Symptom checkers (also called “symptom assessment applications,” “online symptom checkers,” or “self-assessment applications”) are websites or mobile apps in which medical laypeople can enter their symptoms. The apps then provide potential diagnoses and “self-triage” advice. Self-triage advice refers to recommendations given in a precare setting to assist users in determining if, how urgently, and in which institution they should seek care. The first study to systematically analyze the accuracy of these apps was conducted in 2015, and their accuracy has been debated ever since [1]. This seminal study evaluated symptom checkers using 45 medical case vignettes (15 emergency care cases in which users would call the national emergency line or go directly to the emergency department, 15 nonemergency cases in which users would seek primary care, and 15 self-care cases in which users would treat symptoms themselves or wait to see if symptoms improve before seeking care) that were taken from various medical resources, including medical education textbooks. The gold standard solution—that is, the most appropriate action for each case—was determined by 2 physicians who independently rated each case and then discussed disagreements. An unrelated researcher entered all cases into the various symptom checkers, and the authors calculated the proportion of cases correctly solved as the main outcome. This procedure has been used in most subsequent studies, sometimes with slight modifications such as adding more vignettes and triage levels, using lay-friendly phrasing of the vignettes, or including large language models as symptom checkers [2-5]. However, most of these studies acknowledged limitations with this approach and called for improved methods. Systematic reviews that attempted to determine the accuracy of symptom checkers across multiple studies quickly reached the consensus that these methods were often of low quality and that cross-study comparability was limited [6-9]. In recent years, some studies have explicitly formalized this criticism, whereas others have proposed solutions to address it, including several of our own [6,10-14].

In this paper, we do not want to add to this criticism; instead, we present the Symptom Checker Accuracy Reporting Framework (SCARF) and a checklist that can (1) be used to conduct high-quality symptom checker evaluation studies and (2) standardize the evaluation procedure to increase cross-study comparability of symptom checkers and large language models. Because self-triage advice is arguably the most useful information for medical laypeople, this framework focuses on self-triage accuracy as the main outcome [15].

Limitations and Challenges of Previous Methodologies

Most studies evaluating the triage accuracy of symptom checkers have criticized the existing evaluation approach for being artificial. In particular, the vignettes describe idealized, unambiguous cases, and some include scenarios for which symptom checkers would rarely be consulted (eg, recurrent aphthous stomatitis, which may be unexplainable upon first appearance but is easily recognized once experienced [16]). If the aim is to determine a triage accuracy metric that can be generalized to real-world interactions and scenarios in which symptom checkers are actually used, the inclusion of such cases in evaluations seems questionable. Apart from vignettes, current evaluation approaches have several other shortcomings that we grouped into 4 categories: generalizability, symptom input, gold standard assignment, and metrics (Figure 1). We build on these points to develop our standardized methodological evaluation framework.

Figure 1. Four categories of criticism regarding symptom checker evaluation studies.



The first point concerns the generalizability of the evaluations. This includes both the vignettes and the experimental setup, which, according to ecological validity theory, should resemble real-world use cases and interactions to yield results that can be generalized [17]. Traditional vignettes have been criticized for a lack of representativeness for several reasons. First, they are often derived from medical education textbooks and are therefore artificial, not representing the unspecific concerns for which patients would use a symptom checker [10,14]. Second, these cases are mostly written post hoc by clinicians who have access to more specialized information (eg, diagnoses, laboratory test results, and clinical examinations) than a patient consulting a symptom checker [10,13,18]. In other words, thus far, the information in existing vignettes does not reflect the types of problems actual users of symptom checkers face, and it is not clear what that information should be. Third, the cases described in the vignettes do not reflect the natural base rates of emergency or nonemergency versus self-care cases among users [10,13,14]. Fourth, there is no consensus on the number of vignettes that should be included in a vignette set or how to ensure their quality [10]. The experimental setup focuses on symptom checker accuracy and has thus been criticized for implicitly assuming that symptom checker advice directly translates into user actions, even though symptom checkers merely provide advice that users may or may not follow [19-21]. This limitation confines research to assessing only the "technical accuracy" of a symptom checker, without addressing its likely real-world impact. To determine whether technical accuracy translates into improved decision-making by users, symptom checkers ultimately need to be evaluated in user studies.

The second point concerns the procedure for inputting symptoms. Typically, a single person—who may or may not have medical expertise—enters the symptoms. Because not all information that a symptom checker might ask for is included in the vignette, the inputter must make assumptions about the case when asked about it. Thus, clinicians tend to rely on their clinical judgment and expertise, whereas laypeople—the actual users of symptom checkers—use various strategies, ranging from guessing to ignoring the questions they are asked [10,11,22]. It is also unclear how many inputters should be involved in the evaluation to yield valid performance estimates [10]. These issues suggest that the final output is highly dependent on the inputter or inputters, which creates an information bias that limits the internal validity of evaluation studies. This information bias is further compounded by the fact that different symptom checkers allow different input modalities (eg, free text, multiple-choice questions, images, or even laboratory results). These variations introduce an inherent comparability limitation, as the same case may be assessed differently depending on the input options of the tool.

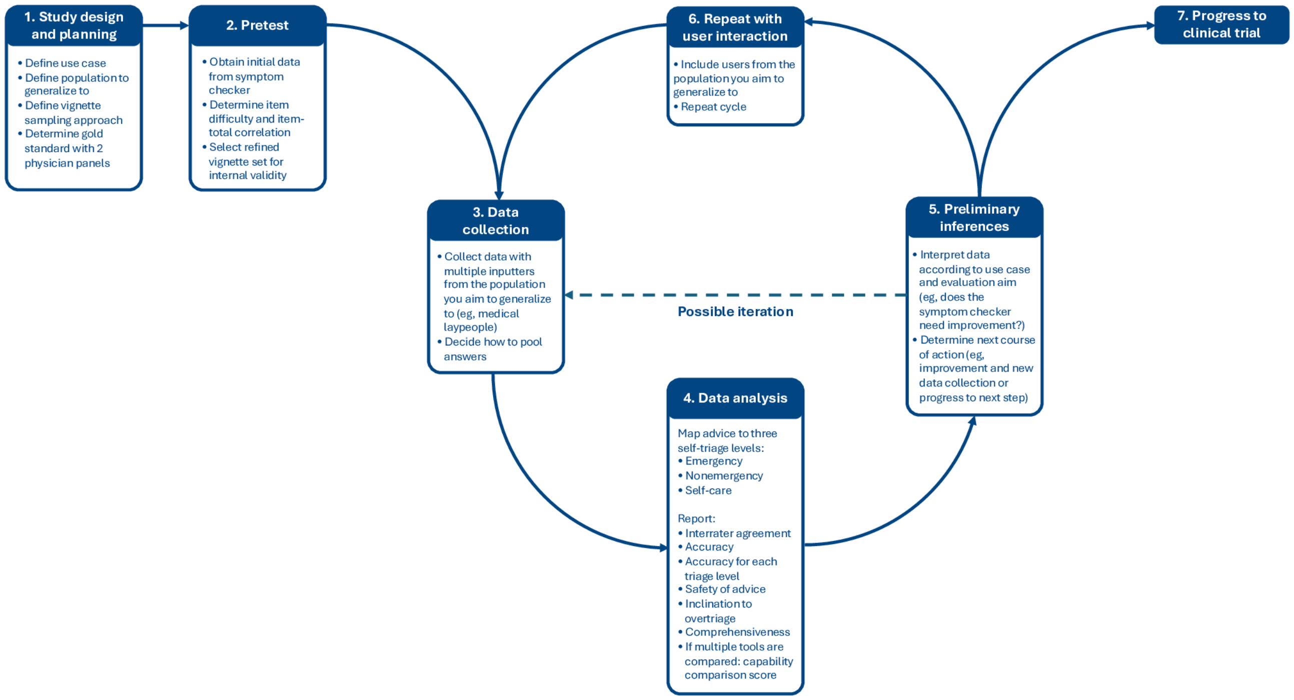
The third point relates to the gold standard assignment used to assign the solution to a case vignette. Different studies use varying procedures: some use a single physician, others use multiple physicians, recordings from clinical encounters (such as telephone triage), or sometimes the authors even determine the gold standard solution themselves [10,12]. This variation not only limits the comparability between studies but also raises questions about the accuracy of the assigned gold standard in some cases.

The final point concerns the metrics used to evaluate symptom checkers. Although most studies report triage accuracy as the proportion of cases solved correctly, some also include additional metrics, such as the tendency to overtriage or undertriage and the safety of the advice [14]. The exact self-triage classifications differ between studies as well: for example, Semigran et al [1] used a 3-tiered approach including "emergencies," "nonemergencies," and "self-care cases," whereas Hill et al [2] extended this classification to include "1-day urgent cases." Furthermore, because different symptom checkers use varying classifications as well, it is unclear how their advice should be mapped to the study's triage categories (eg, whether an urgent care clinic is considered emergency or nonemergency care). These issues ultimately limit cross-study comparability.

Framework

To address these points, we developed an evaluation framework by integrating available empirical studies on methodological improvements. This framework can be found in Figure 2. It can be used for preclinical evaluations to identify symptom checkers that are likely to perform well in clinical trials and real-world evaluations. Once identified, the symptom checker should nonetheless undergo testing in a 3-phase clinical trial similar to pharmaceutical trials [23]. Hence, our framework not only standardizes vignette-based symptom checker evaluations but also makes subsequent clinical trials more cost efficient by identifying tools likely to yield positive outcomes. It can be applied both to evaluations across a broad set of cases as well as to those focusing on specific patient groups (eg, patients receiving rheumatology care), by defining the intended use case and population accordingly.

Figure 2. Integrated preclinical evaluation framework (Symptom Checker Accuracy Reporting Framework; SCARF) for evaluating the self-triage accuracy of symptom checkers and artificial intelligence–based tools.



In the beginning (part 1), evaluators should clearly define the use case they intend to examine, such as “self-triage decisions” or “emergency care decisions.” Next, they should specify the target population to which they wish to generalize. For the self-triage use case, this might include symptom checker users deciding on their next course of action. Then, they should define a vignette sampling approach, which ensures that vignettes are representative of real patient cases and accurately reflect the disease or symptom and triage prevalence relevant to the use case. For example, the approach could sample real patient cases stratified according to the prevalence of symptom types entered into symptom checkers. A systematic sampling procedure to do that is available in the RepVig framework, and for the self-triage use case, a representative vignette set is provided in the framework’s validation study [13]. At this stage, researchers should also assign a gold standard solution to each case and define how the possible outputs of a symptom checker are mapped onto these categories. According to a study by El-Osta et al [12], this should involve 2 physician panels that independently rate the cases in focus groups and resolve any disagreement through discussion until consensus is reached.

Next, evaluators should obtain initial data from some symptom checkers to refine the vignette set according to test-theoretical criteria (part 2) to ensure that vignettes are not only externally but also internally valid. This process involves calculating the item-total correlation and excluding any cases with a negative or zero item-total correlation (to ensure that only cases accurately predicting overall performance are included). Additionally, item difficulty for each vignette should be determined, and cases with an item difficulty of zero may be excluded (to ensure that vignettes add meaningful information and are not impossible to solve). However, if these cases can be solved by physicians and are clinically plausible, even items with an item difficulty of zero may be retained in the vignette set to avoid inflating performance estimates. A procedure for this is outlined in one of our previous studies [24]. The size of the final vignette set should ultimately be determined using a power analysis. However, given that entering a large number of vignettes manually may be infeasible and that there is no empirical evidence on optimal set sizes, we pragmatically recommend a minimum of 45 vignettes. This number has proven feasible and has been applied across multiple studies [1,2,13,25], as it can be developed and entered by a single evaluator within a reasonable time frame, while still providing sufficient variation for a statistical analysis.

Using the refined vignette set, data from all symptom checkers can be collected (part 3). Multiple inputters (at least 2, possibly more) should enter every case into each symptom checker and select a “not sure” option in cases of missing information. To minimize inputter variability, inputters should follow a standardized protocol. For instance, Meczner et al [11] instructed inputters to enter only the symptoms explicitly stated in the vignette, allowing synonyms or broader categories but rejecting new information not included in the vignette. Their publication provides entry instructions that can be used in future studies to standardize input procedures. Once the inputters have obtained the data, their results should then be pooled. This can be achieved using several algorithms, but the best approach appears to be a majority vote, that is, the advice most frequently given to all evaluators [11]. For example, if 2 inputters receive the advice to seek emergency care while 1 inputter receives self-care advice, the recommendation should be coded as “emergency”.

In the next step, the data analysis (part 4), evaluators should map the received advice to a multitiered classification system. To increase comparability across studies and health care systems, we suggest using a 3-tiered classification system—“emergency,” “nonemergency,” and “self-care”—to provide a common reference structure. A potential

"1-day-urgent" category could be classified as "nonemergency." At the same time, we acknowledge that some systems use more granular triage categories; therefore, we suggest conducting sensitivity analyses (eg, treating "1-day-urgent" as "emergency" or as its own category, or analyzing the full set of available tiers) to assess the stability of the results. After mapping each recommendation, evaluators should first report the interrater reliability among all inputters to identify the influence of different inputters, followed by a set of metrics: overall accuracy, accuracy for each triage level, safety of advice, inclination to overtriage, and comprehensiveness [14,26]. These metrics were identified through systematic review of previously reported metrics and can increase comparability across different studies [14,26]. If multiple symptom checkers are evaluated simultaneously, we propose additionally reporting the Capability Comparison Score (developed in a previous study) to determine how symptom checkers perform relative to each other [14,26]. To assist researchers in reporting and visualizing these metrics, the R package *symptomcheckR* is available, where the formulas for calculating all metrics are described as well [26].

In the next step, the results should be interpreted according to the defined use case, and the next course of action should be determined (part 5): if developers aim to validate their tools, they may either decide to improve their tool and test it again (by going back to step 3) using the same setting or continue with the evaluation and test the best-performing tools with users in the loop making self-triage decisions (step 6). In this phase, users should be provided with the symptom checker, and the tool's impact, instead of its "technical accuracy," should be assessed in a new evaluation with sufficient statistical power [21]. This step is included because preclinical vignette studies can only benchmark technical accuracy and do not capture whether laypeople actually make better decisions when using a symptom checker. If results of user studies are also promising, the symptom checker can then be tested in a clinical trial with real patients and their symptoms to assess whether the symptom checker advice also translates to improved decisions by users in the real world (step 7).

Open Questions

Our approach leaves several open questions for future research. First, some of the vignettes (such as the vignettes by Semigran et al [1] and our own [13]) do not include additional information for questions that symptom checkers may ask. Although some vignette sets do include additional information, there is no universal way to collect additional information. Future research could develop a method to supplement this missing information—perhaps using a hybrid approach that combines interviews with patients from whom the case vignettes were derived and synthetic artificial intelligence (AI)–generated supplementary data based on these interviews. Second, it remains unclear whether "accuracy" or a "correct" solution should be the main outcome. Perhaps a binary classification of correct versus incorrect in a task like symptom assessment that is associated with high uncertainty may be less relevant than assessing the impact of the advice—specifically, whether it is safe and appropriate for the individual and whether it increases or decreases health care demands. Third, with the introduction of large language models as an alternative to traditional symptom checkers, output variability plays an even greater role. Future research should address how to manage the variability of generated outputs when provided with identical inputs. Fourth, current evaluations do not specifically include atypical presentations. It remains unclear whether case vignettes are only suitable for typical cases or if vignette sets for atypical cases could also be developed. Although the RepVig framework could be used for developing such a vignette set again, assigning a reliable gold standard solution to atypical cases will be challenging [13]. Finally, our approach is highly tailored to a self-triage use case. Although it standardizes most aspects of an evaluation, diagnostic use cases may require additional details (such as clinical plausibility of the vignettes or a procedure to determine whether a diagnosis matches the gold standard) and outcome metrics (such as cumulative diagnostic scores [27]) that are not covered by our approach.

Outlook

The SCARF (and the corresponding checklist in Multimedia Appendix 1 and Multimedia Appendix 2) presented in this paper addresses all previously raised points of criticism and aims to improve the quality and comparability of future symptom checker evaluations. However, we acknowledge that the presented approach is more resource intensive than the traditional approach introduced by Semigran et al [1] and may not be feasible for every evaluation. To aid researchers in integrating these methods into practice, several open resources are available for the presented use case: for example, representative vignettes are openly accessible and free to use [13], a refined vignette set that satisfies test-theoretical criteria is available as well [24], and all metrics can be easily calculated using the open-source symptomcheckR package [26]. We encourage researchers to build on these resources to improve the quality of future evaluations and enhance cross-study comparability.

Conclusions

In this paper, we summarize the limitations and challenges of previous studies evaluating symptom checkers using vignettes. In recent years, several empirical studies have addressed most of these limitations individually, yet a unified methodological and reporting framework integrating these findings was missing. We present a preclinical framework and the corresponding SCARF checklist upon which future vignette-based symptom checker evaluations can build to address generalizability, input variability, gold standard assignment, and metrics, and we highlight several open access resources that evaluators can use. By adopting this approach, researchers can identify well-performing tools for more cost-efficient clinical trials and can significantly increase the quality and comparability of vignette-based symptom checker evaluation studies, thereby enabling reliable evidence syntheses. This can help move closer to assessing and improving the effectiveness of symptom checkers, diagnostic decision support systems, and large language models.

Funding

The authors acknowledge support from the Open Access Publication Fund of Technische Universität Berlin.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Symptom Checker Accuracy Reporting Framework (SCARF) checklist (editable version).
[DOCX File, 22 KB - humanfactors_v13i1e76168_app1.docx]

Multimedia Appendix 2
Symptom Checker Accuracy Reporting Framework (SCARF) checklist (PDF version).
[PDF File, 82 KB - humanfactors_v13i1e76168_app2.pdf]

References

1. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. BMJ 2015 Jul 8;351(1–9):h3480. [doi: 10.1136/bmj.h3480] [Medline: 26157077]
2. Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. Med J Aust 2020 Jun;212(11):514-519. [doi: 10.5694/mja2.50600] [Medline: 32391611]
3. Ceney A, Tolond S, Glowinski A, Marks B, Swift S, Palser T. Accuracy of online symptom checkers and the potential impact on service utilisation. PLoS ONE 2021;16(7):e0254088. [doi: 10.1371/journal.pone.0254088] [Medline: 34265845]
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. [doi: 10.2196/31810] [Medline: 35536633]
5. Ito N, Kadomatsu S, Fujisawa M, et al. The accuracy and potential racial and ethnic biases of GPT-4 in the diagnosis and triage of health conditions: evaluation study. JMIR Med Educ 2023 Nov 2;9:e47532. [doi: 10.2196/47532] [Medline: 37917120]
6. Wallace W, Chan C, Chidambaram S, et al. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. NPJ Digit Med 2022 Aug 17;5(1):118. [doi: 10.1038/s41746-022-00667-w] [Medline: 35977992]
7. Chambers D, Cantrell AJ, Johnson M, et al. Digital and online symptom checkers and health assessment/triage services for urgent health problems: systematic review. BMJ Open 2019 Aug 1;9(8):e027743. [doi: 10.1136/bmjopen-2018-027743] [Medline: 31375610]
8. Riboli-Sasco E, El-Osta A, Alaa A, et al. Triage and diagnostic accuracy of online symptom checkers: systematic review. J Med Internet Res 2023 Jun 2;25:e43803. [doi: 10.2196/43803] [Medline: 37266983]
9. Kopka M, von Kalckreuth N, Feufel MA. Accuracy of online symptom assessment applications, large language models, and laypeople for self-triage decisions. NPJ Digit Med 2025 Mar 25;8(1):178. [doi: 10.1038/s41746-025-01566-6] [Medline: 40133390]
10. Painter A, Hayhoe B, Riboli-Sasco E, El-Osta A. Online symptom checkers: recommendations for a vignette-based clinical evaluation standard. J Med Internet Res 2022 Oct 26;24(10):e37408. [doi: 10.2196/37408] [Medline: 36287594]
11. Meczner A, Cohen N, Qureshi A, et al. Controlling inputter variability in vignette studies assessing web-based symptom checkers: evaluation of current practice and recommendations for isolated accuracy metrics. JMIR Form Res 2024 May 31;8:e49907. [doi: 10.2196/49907] [Medline: 38820578]
12. El-Osta A, Webber I, Alaa A, et al. What is the suitability of clinical vignettes in benchmarking the performance of online symptom checkers? An audit study. BMJ Open 2022 Apr 27;12(4):e053566. [doi: 10.1136/bmjopen-2021-053566] [Medline: 35477872]
13. Kopka M, Napierala H, Privoznik M, Sapunova D, Zhang S, Feufel MA. The RepVig framework for designing use-case specific representative vignettes and evaluating triage accuracy of laypeople and symptom assessment applications. Sci Rep 2024 Dec 23;14(1):30614. [doi: 10.1038/s41598-024-83844-z] [Medline: 39715767]
14. Kopka M, Feufel MA, Berner ES, Schmieding ML. How suitable are clinical vignettes for the evaluation of symptom checker apps? A test theoretical perspective. Digit Health 2023;9:20552076231194929. [doi: 10.1177/20552076231194929] [Medline: 37614591]
15. Aboueid S, Meyer S, Wallace JR, Mahajan S, Chaurasia A. Young adults' perspectives on the use of symptom checkers for self-triage and self-diagnosis: qualitative study. JMIR Public Health Surveill 2021 Jan 6;7(1):e22637. [doi: 10.2196/22637] [Medline: 33404515]
16. Baccaglini L, Theriaque DW, Shuster JJ, Serrano G, Lalla RV. Validation of anamnestic diagnostic criteria for recurrent aphthous stomatitis. J Oral Pathology Medicine 2013 Apr;42(4):290-294. [doi: 10.1111/jop.12015]

17. Brunswik E. Representative design and probabilistic theory in a functional psychology. Psychol Rev 1955 May;62(3):193-217. [doi: 10.1037/h0047470] [Medline: 14371898]
18. Yu SWY, Ma A, Tsang VHM, Chung LSW, Leung SC, Leung LP. Triage accuracy of online symptom checkers for accident and emergency department patients. Hong Kong J Emerg Med 2020 Jul;27(4):217-222. [doi: 10.1177/1024907919842486]
19. Verzantvoort NCM, Teunis T, Verheij TJM, van der Velden AW. Self-triage for acute primary care via a smartphone application: practical, safe and efficient? PLoS ONE 2018;13(6):e0199284. [doi: 10.1371/journal.pone.0199284] [Medline: 29944708]
20. Kopka M, Schmieding ML, Rieger T, Roesler E, Balzer F, Feufel MA. Determinants of laypersons' trust in medical decision aids: randomized controlled trial. JMIR Hum Factors 2022 May 3;9(2):e35219. [doi: 10.2196/35219] [Medline: 35503248]
21. Kopka M, Wang SM, Kunz S, Schmid C, Feufel MA. Technology-supported self-triage decision making. NPJ Health Syst 2025 Jan 25;2(1):1-11. [doi: 10.1038/s44401-024-00008-x]
22. Fraser H, Coiera E, Wong D. Safety of patient-facing digital symptom checkers. The Lancet 2018 Nov;392(10161):2263-2264. [doi: 10.1016/S0140-6736(18)32819-8] [Medline: 30413281]
23. You JG, Hernandez-Boussard T, Pfeffer MA, Landman A, Mishuris RG. Clinical trials informed framework for real world clinical implementation and deployment of artificial intelligence applications. NPJ Digit Med 2025 Feb 17;8(1):107. [doi: 10.1038/s41746-025-01506-4] [Medline: 39962232]
24. Kopka M, Feufel MA. Statistical refinement of patient-centered case vignettes for digital health research. Front Digit Health 2024;6:1411924. [doi: 10.3389/fdgth.2024.1411924] [Medline: 39498100]
25. Levine DM, Tuwani R, Kompa B, et al. The diagnostic and triage accuracy of the GPT-3 artificial intelligence model: an observational study. Lancet Digit Health 2024 Aug;6(8):e555-e561. [doi: 10.1016/S2589-7500(24)00097-9] [Medline: 39059888]
26. Kopka M, Feufel MA. Software symptomcheckR: an R package for analyzing and visualizing symptom checker triage performance. BMC Digit Health 2024 Jul 22;2(1):43. [doi: 10.1186/s44247-024-00096-7]
27. Knitza J, Hasanaj R, Beyer J, et al. Comparison of two symptom checkers (ada and symptoma) in the emergency department: randomized, crossover, head-to-head, double-blinded study. J Med Internet Res 2024 Aug 20;26:e56514. [doi: 10.2196/56514] [Medline: 39163594]

Abbreviations

AI: artificial intelligence
SCARF: symptom checker accuracy reporting framework

Edited by A Kushniruk; submitted 17.Apr.2025; peer-reviewed by A Meczner, AJ Wetzel, J Knitza; accepted 20.Nov.2025; published 16.Jan.2026.

Please cite as:
Kopka M, Feufel MA
How to Evaluate the Accuracy of Symptom Checkers and Diagnostic Decision Support Systems: Symptom Checker Accuracy Reporting Framework (SCARF)
JMIR Hum Factors 2026;13:e76168
URL: https://humanfactors.jmir.org/2026/1/e76168
doi:10.2196/76168

Expert Views on Criteria for Evaluation of Human Factors Methods: Qualitative Interview Study

Selvana Awad[1,2], BPharm, MHSM; Rachel Begg[2], BPsych, MPsych; Thomas Loveday[2], BPsych, MPsych, PhD; Andrew Baillie[1,3], BSc, MPsych, PhD; Melissa Baysari[1], BPsych, PhD

[1]The University of Sydney, Camperdown, Sydney, Australia
[2]NSW Health, 1 Reserve Rd, St Leonards, Australia
[3]Sydney Local Health District, Camperdown, Australia

Corresponding Author:
Selvana Awad, BPharm, MHSM
The University of Sydney, Camperdown, Sydney, Australia

Abstract

Background: Human factors (HF), or ergonomics, which explores the interaction between humans and systems, has been used to support design in safety-critical industries such as aviation, transportation, nuclear power, and manufacturing. HF methods have the potential to support the safe design of health IT; however, the evaluation of HF methods to determine their effectiveness and feasibility in this context has been limited.

Objective: The aim of this study was to identify criteria for evaluating HF methods when applied to real-world projects and to use these to propose a framework for method evaluation.

Methods: The study design was qualitative and descriptive and involved semistructured interviews with HF experts working across health and nonhealth industries in academic and/or practitioner roles. HF experts held a relevant degree (eg, ergonomics and HF engineering) and were actively using their HF expertise. Results were thematically analyzed.

Results: A total of 21 participants took part, and interviews lasted, on average, 52 (range 39 - 103) minutes. Participants mentioned that they did not routinely evaluate methods; however, when asked how they would evaluate methods, they outlined a range of criteria to support method evaluation. Overall, 5 criteria and 28 subcriteria were identified. High-level criteria included effectiveness, efficiency, ease of use and acceptability, and impact on the solution.

Conclusions: Results from this study were used to propose a framework for evaluating HF methods used in real-world health IT projects. The framework should provide organizations with valuable information on how to optimize the application and outcomes of HF methods and build HF capability within organizations, particularly where this capability may be lacking.

(JMIR Hum Factors 2026;13:e73324) doi:10.2196/73324

KEYWORDS

human factors; evaluation; methods; quantitative; qualitative; framework

Introduction

Health IT (HIT) enables the processing, storage, and exchange of health information in an electronic environment [1]. In the United States, the Meaningful Use program has led to a significant increase in the adoption of HIT, such as electronic health records [1]. Although clear benefits of HIT have been demonstrated [2,3], research has shown that many HITs suffer from poor usability [4,5]. Usability is a measure of how well a specific user in a specific context can use a product to achieve a defined goal safely, effectively, efficiently, and satisfactorily [5]. Usability can be compromised by poor design and lead to “use errors” [5-7]. The relationship between system design and safety in the context of HIT is well recognized in the literature [5,8-12]. For example, a recent systematic review found that poorly designed electronic health records were associated with usability issues such as poor data entry, lack of workflow support, and inadequate automation [13]. These issues directly contributed to medication errors, such as patient overdoses, and had other negative impacts on medication safety [13]. Human factors (HF) methods have the potential to reduce “use errors” and, in turn, improve patient outcomes [7].

The discipline of HF, or ergonomics, which explores the interaction between humans and systems, has been used to support design in safety-critical industries such as aviation, transportation, nuclear power, and manufacturing [7,14-16]. While some HF methods, particularly those focused on human-computer interaction (eg, usability testing), have been used to support the design and redesign of HIT, the use of safety-focused and systems-based HF methods in this context is limited [8,17]. HF methods commonly used in the HIT context tend to have a linear and micro perspective of potential issues that may affect system use and user behavior, as they focus on specific problems that may be encountered by the individual user using the system rather than considering the entire sociotechnical system [8,17,18]. HF methods that apply a systems thinking lens to identify problems and risks that may arise from the interactions between components of a complex system are less commonly used in HIT design and evaluation [8,17,19,20].

Potential reasons for the limited use of HF methods include issues with the availability of HF expertise in health care, a research-practice gap, a potential echo chamber effect within health care whereby the same subset of well-known methods is used repeatedly, issues with method usability, and challenges with demonstrating the value added by HF application to justify up-front investment [10,17,20,21]. Further research is required to demonstrate the value of applying HF methods (particularly those with a systems safety focus) to HIT, supported by a robust evaluation approach or framework [22,23].

Evaluation of HF methods to determine their effectiveness and feasibility has been limited. Previous evaluations have mainly been in the context of academic studies evaluating the reliability, validity, and efficacy of some HF methods, for example, when applied by health care participants as compared with experienced HF experts [20]. Despite these studies, empirical data regarding the reliability or validity of many HF methods do not exist [24]. Several challenges associated with conducting these forms of evaluations have been cited, including challenges with recruiting enough experts to enable comparison with a gold standard, time and resource intensiveness, and limited knowledge of appropriate statistical analyses [22,24].

An additional challenge is that there is no consensus or agreement on what constitutes an effective or valuable HF method. How do HF practitioners select and evaluate methods when applying them to real-world projects as part of HF integration processes?

Although a previous paper provides suggestions on potential evaluation criteria for HF methods [18], these are not comprehensive and represent the authors' recommendations rather than findings derived from research methods. The aim of this study was to identify criteria for evaluating HF methods when applied to real-world projects and to use these to propose a framework for method evaluation.

Methods

Overview

This qualitative descriptive study was undertaken as part of a larger study that focused on HF and safety analysis methods for use in the design, redesign, and configuration of HIT. The study design involved semistructured interviews with HF experts working across health and nonhealth industries in academic and/or practitioner roles. A participant was considered an HF expert if they had a relevant degree (eg, ergonomics and HF engineering) and were actively using their HF expertise. Part 1, currently under review, explored what HF methods experts use and how they are selected. Part 2, reported here, explored criteria HF experts use or view as important to evaluate HF methods. Questions used to guide the semistructured interviews were developed by a clinical informatics professional with expertise in design, HF, and safety and quality (SA) and a HF expert (MB) and reviewed by other members of the research team with HF (TL and RB) and implementation science (AB) expertise. After collecting demographic information, the 2 main questions asked relevant to this study were "What makes a 'good' HF method to you?" and "How would you evaluate a HF method?"

Participants were recruited through a combination of purposive, opportunistic, and snowball sampling. Recruitment involved advertising the study through national HF societies and relevant working groups; promotion at an international HF and patient safety conference through networking and word-of-mouth approaches; consultation with HF experts and contacts to recommend potential participants; and a review of common HF textbooks and literature to identify HF authors who could be invited to participate. In addition, HF practitioners known to the research team were directly approached. Suitable participants interested in the study, including those identified by the investigators, were invited to take part in the study via email. Recruitment continued until thematic saturation was reached, that is, no new themes were emerging from the data [25].

The interviews, conducted by SA, occurred via videoconferencing, except for one, which was face to face. Deidentified content from the interview transcripts was independently analyzed by 2 investigators with expertise in HF and qualitative analysis (SA and RB) using a general inductive approach [26]. Each investigator independently coded data for the first 5 interviews and then met to discuss findings and reach consensus on a high-level framework to support the documentation of codes into themes and subthemes. For the remainder of the interviews, each investigator continued to independently assign text to the agreed themes and subthemes using the framework and, at the end of this process, met to discuss any further discrepancies until consensus was reached [27]. Overall, the 2 investigators were generally consistent in their coding and identification of themes and subthemes. Disagreements were minor and resolved via discussion until consensus was reached.

Ethical Considerations

Ethics approval was obtained from the University of Sydney's Human Research Ethics Committee. All participants provided informed consent and agreed to be audio-recorded. All transcripts were de-identified prior to data analysis. All participants provided informed consent and agreed to be audio-recorded.

Results

Participant Demographics

A total of 21 participants took part, and interviews lasted, on average, 52 (range 39 - 103) minutes. Table 1 describes the demographics of the 21 participants.

Table . Characteristics of the participants included in the interviews (N=21).

Core industry and core role		Participants, n	Within industry category, %	Within total participants, %
Health (n=14)				
	Academic	9	64	43
	Practitioner	2	14	10
	Academic and practitioner	3	21	14
	Total	14	100	67
Nonhealth (n=7)				
	Academic	2	29	10
	Practitioner	5	71	24
	Academic and practitioner	0	0	0
	Total	7	100	33

How Should HF Methods Be Evaluated?

Participants explained that they did not routinely evaluate methods; however, when asked how they would evaluate them, they outlined a range of criteria to support method evaluation (Table 2). High-level criteria included effectiveness, efficiency, ease of use and acceptability, and impact on the solution. Other criteria, such as validity, reproducibility, and reliability, were mainly mentioned by academic participants rather than practitioners. Overall, 5 criteria and 28 subcriteria were identified.

Table . Criteria for human factors (HF) method evaluation, as reported by participants.

Criteria and subcriteria	Example quotes
Effectiveness	
1.1 Effectiveness in identifying usability issues	• [N]ice to see whether or not the amount of usability issues related to patient safety. [P2]
1.2 Effectiveness in identifying safety issues	• [N]ice to see whether or not the amount of usability issues related to patient safety. [P2] • ...there's a primary metric that matters most like which method uncovered the most safety issues in advance of implementation or which method uncovered the more severe safety issues [P9]
1.3 Ability to achieve intended impact, change or goal	• What value did it add? Like? Do you feel like added value to the project? Do you feel like it gave you the right answers or the right tools to get to the answers that you wanted? [P16] • Does it make the change that you envision?...And that has to do with your final goal [P6]
1.4 Generates recommendations that are useful and easy to implement	• Does it generate recommendations [that] are easy to implement? And that then do [they] get implemented and lead to risk management? [P7] • [T]he other thing that's really important to us in terms of criteria is going beyond analysis and figuring out a method that can actually produce useful information, and actionable information for redesign for change, for implementation, whatever you want [P11]
1.5 Ability to identify micro, meso and macro (systems) level considerations that have safety implications	• [Y]ou might do an overarching review of the technology in the context of the people using it in the context of the work in the workplace and then on the basis of that, deploy specific things that are the micro, meso, or macro level, to understand a bit more [P7]
1.6 Method's effectiveness in understanding the dynamic nature of the system	• And another important thing is to continue to collect data about the use of the system, what I call the dynamic safety of an application. [P13]
1.7 Ability to cover the required domain areas/constructs (eg, usability, safety, workload, situational awareness, and decision-making)	• And then I think it's the core human factors constructs. It's kind of like, what's the impact on situational awareness, workload? Usability, and decision making, are probably some key constructs that you might want to check [P21]
1.8 Overall usefulness	• [T]he other thing that's really important to us in terms of criteria is going beyond analysis and figuring out a method that can actually produce useful information, and actionable information for redesign for change, for implementation, whatever you want [P11]
Efficiency	
2.1 Demand on time and resources (e.g.eg, workload)	• How fast it was... how much time [and] how much budget it came under. [P16]
2.2 Cost required	• Is it cost effective? [P6]
2.3 Overall efficiency	• It's to look at the efficiency of the methods [P4]
Ease of use and acceptability	
3.1 Ease of use	• Is it usable, based on some objective standard usability? Or is it more usable than the alternative? [P7] • [Y]our method should be easy to use... not too demanding in human resources in training, and not too costly [P1] • [E]ase of use of the method and the ease of understanding the tools from the people who participate [P13]
3.2 Learnability	• How quick it was to learn [P16]

Criteria and subcriteria	Example quotes
3.3 Utility and acceptability of the method	• The types of questions that make sense are acceptability. You know, utility...in [a] very basic sense, we applied this in this way, what did we learn, we learned something new [P7]
3.4 Overall satisfaction of those applying the method	• [E]ase of use of the method and the ease of understanding the tools from the people who participate [P13]
3.5 Participant experience	• [E]ase of use of the method and the ease of understanding the tools from the people who participate [P13] • [W]e asked participants to give us feedback on the method...I think that's an important source of feedback and evaluation of the method, what do they think about it? [P11]
3.6 Likely adoption of the method based on complexity and learnability	• If you develop a very complex method, it may be the best one to identify all the problems that could lead to a risk for the patient. But...nobody's applying it [P1]
3.7 Adoption of the method by non-HF experts	• [T]alked about giving away ergonomics. And in a lot of the things that we've done that has been our goal when we leave, can they actually do things on their own? Even if it's a little bit more simple than the way we would do it? [P11]
3.8 Whether the method met expectations	• And another important thing probably is to clarify, the time needed to apply one method and the expected outcome, like a sort of table of the expectation that one can have through the application of that method, because too often, people look at the ergonomist, either those with the silver bullet in the hand, or those with some annoying requirements to be applied in the design process. [P13]
3.9 Ability of the method to be adapted	• [D]evelop questions that can be added into evaluation about the potential adaptation of these methods [P7]
Impact	
4.1 Impact on patient safety and other outcomes	• [If] you've got a goal of patient safety...the main criterion should be the risk to the patient... the likelihood of patient safety issues...if you want that your method is used, if you want to make it to be used by vendors and hospital later...analyze their human factors or risk, relate the human factors related risks of EHR [electronic health record] [P1] • Is the whole system in total more safe? [P6]
4.2 Impact on end user workflows and workloads	• [Y]ou can also identify workload issues that will slow down the care process even if it doesn't threaten directly patient [P1]
4.3 End user satisfaction (of the health information technology)	• The satisfaction of the end users also could be analyzed [P1]
4.4 Adoption of the system or tool by end users	• Is it used? Is it actually used? By all your users? [P6]
4.5 Impact on redesign (ie, whether changes or improvements were made)	• So when we redesigned the document, how many of those many, many vulnerabilities or problems were identified in the document? How many of those were we able to address? So, you know, having concrete impact on design, I think is a big issue. [P11]
Other	
5.1 Preestablished validity	• I think these reliability and validity studies are really important [P7]
5.2 Reproducibility of results	• [D]id you align to see if you meet that standard? ...and then do all these things that are for measurements, I think it must be reproducible [P9]
5.3 Reliability if applied by different people (interrater reliability)	• I think these reliability and validity studies are really important [P7]

Participants commented on challenges with quantitatively evaluating methods using the criteria identified and suggested a more qualitative, self-assessment–based evaluation; for example:

> *We usually sit down when we've concluded a project as part of the HF team and we will talk about the methods we used and how we feel they performed, whether we had any pitfalls, whether we wished we had something different, or we've done something different.* [P16]

Discussion

This study identified 5 high-level criteria and 28 subcriteria to support the evaluation of HF methods in real-world HIT projects, as reported by HF practitioners across health and nonhealth industries.

Although previous studies have evaluated HF methods by focusing on validity and reliability [23], this study is the first of its kind to identify a broad range of evaluation criteria for HF methods, as recommended by HF practitioners. There is some overlap between our criteria and those developed by Waterson et al [18], which focused on the evaluation of systems-based methods. As such, this study adds to findings generated by Waterson et al [18] by confirming some of the criteria identified in this previous work. This includes outcomes of the method, the method's robustness, the method's usability and support requirements, aspects related to work domains, and aspects related to different levels (ie, individual, team, and organizational) [18].

Our study adds to this by identifying new criteria, particularly those related to operationalized usability, adoption, and other impacts on users (eg, workload); outcome-focused effectiveness, including the link between usability and safety; and more nuanced consideration of system-level considerations (micro, meso, and macro). Furthermore, the findings from our study offer a combined quantitative and qualitative approach to evaluation, with a focus on actionable insights and bridging theory and practice. These points highlight the value of exploring the views of a range of HF experts working across different industries [18].

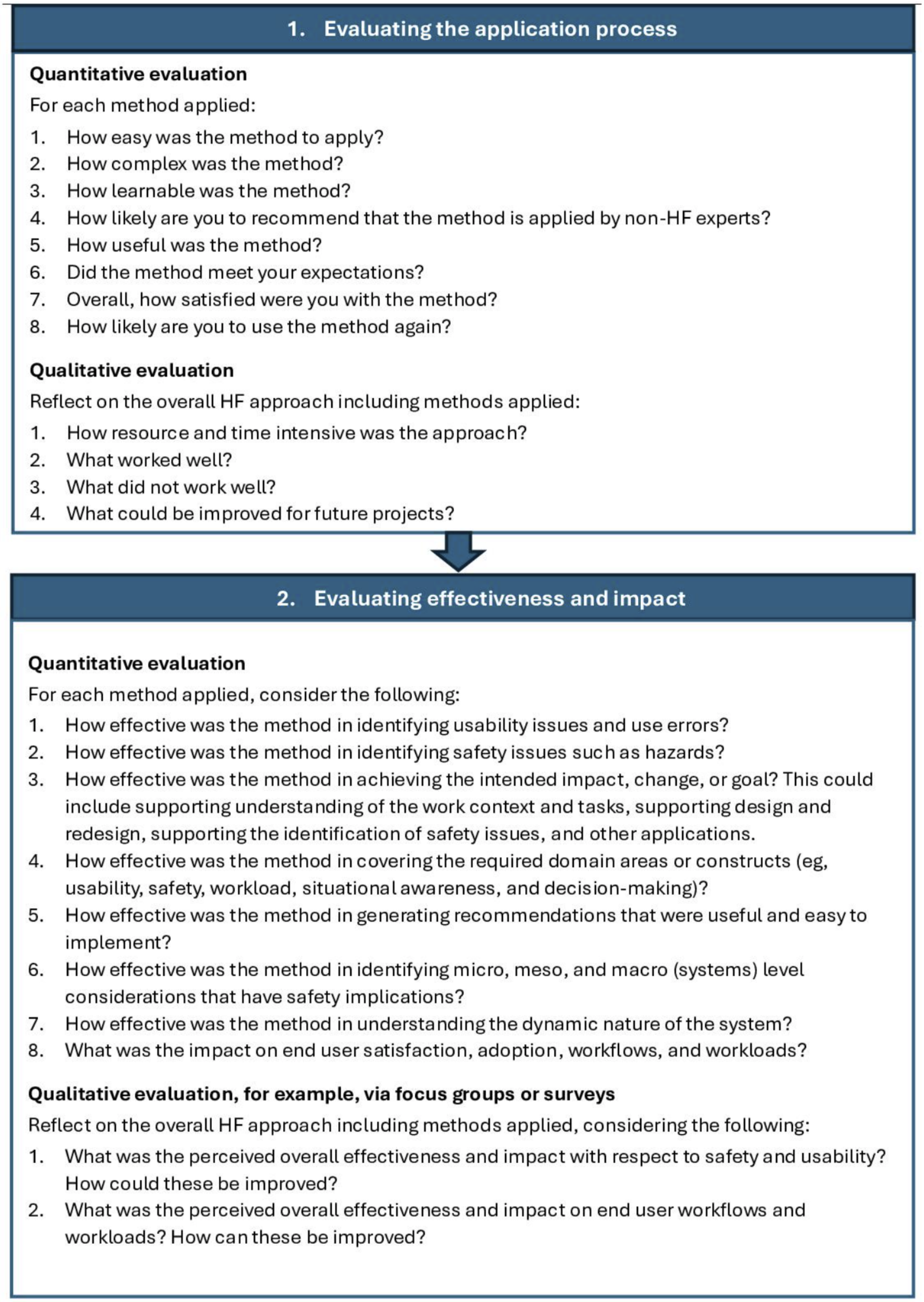
Our study also elaborates on the method's effectiveness in understanding the dynamic nature of the system, which is aligned with systems safety thinking that defines safety as a dynamic, emergent property of how system components interact with each other [19,28-30]. A dynamic system is a complex and adaptive system that changes behavior due to interactions between system components [31-33]. As such, the safety of a system can change over time and is impacted by many variables, such as human performance, resources, and events at particular points in time. This is consistent with a study that aimed to evaluate methods, such as the functional resonance analysis method, using resilience characteristics as indicators of core safety factors [34].

Participants commented on challenges with quantitatively evaluating methods and suggested that a more qualitative approach could be used. This is generally aligned with other literature that suggests that outcome-based quantitative evaluation is not widespread and can be resource intensive [18]. Furthermore, demonstrating the impact of methods can be difficult, as system usability and safety are influenced by a range of factors outside of the method itself [18]. While many factors influence outcomes, health organizations may still benefit from using measures, for example, the number of usability and safety issues identified and user satisfaction scores, to evaluate whether applying methods results in detection of usability and safety issues. Where possible, comparative evaluation to demonstrate that applying methods is better than not applying any method may also be of value. Although such approaches may lack the academic rigor of reliability and validity studies, they may increase an organization's degree of confidence in HF methods and, therefore, willingness to invest in them. Such evaluation may help health staff and managers develop business cases for the application of HF, help organizations reflect on strategies to enhance the use of methods (eg, organizational support), and refine when and how HF methods are applied to deliver the most value within the context of challenging and complex HIT projects. This type of evaluation could occur through self-reported Likert-scale ratings (as done by Waterson et al [18]); reflective qualitative discussions, as suggested by participants in this study; and other relevant metrics.

This study had several limitations. The recruitment process relied on interested experts volunteering to participate, so the sample may have been biased and our results may not represent the views of all HF experts. We included both health and nonhealth participants, as well as practitioners and academics, but did not compare the views of participants from different groups. Although our sample size was modest, we continued interviews until thematic saturation was reached, which is the norm in qualitative research.

We used participant responses and systems safety knowledge to propose a framework for evaluating HF methods in the context of real-world HIT projects (Figure 1). By covering both process- and outcome-based evaluation, this framework loosely aligns with other evaluation frameworks, such as those used in quality improvement and program evaluation [35,36]. While we also recommend further reliability and validity academic studies to validate the robustness of HF methods in the HIT context, our proposed framework offers a subjective yet structured approach for HF method validation that can be applied by organizations to real-world HIT projects where HF methods have been used. The framework, which we recommend be applied flexibly depending on the nature of the project and the resources available, should provide organizations with valuable information on how to optimize the application and outcomes of HF methods and build HF capability within organizations, particularly where this capability may be lacking.

Figure 1. Proposed framework for evaluating human factors (HF) methods.



1. Evaluating the application process

Quantitative evaluation

For each method applied:

1. How easy was the method to apply?
2. How complex was the method?
3. How learnable was the method?
4. How likely are you to recommend that the method is applied by non-HF experts?
5. How useful was the method?
6. Did the method meet your expectations?
7. Overall, how satisfied were you with the method?
8. How likely are you to use the method again?

Qualitative evaluation

Reflect on the overall HF approach including methods applied:

1. How resource and time intensive was the approach?
2. What worked well?
3. What did not work well?
4. What could be improved for future projects?

2. Evaluating effectiveness and impact

Quantitative evaluation

For each method applied, consider the following:

1. How effective was the method in identifying usability issues and use errors?
2. How effective was the method in identifying safety issues such as hazards?
3. How effective was the method in achieving the intended impact, change, or goal? This could include supporting understanding of the work context and tasks, supporting design and redesign, supporting the identification of safety issues, and other applications.
4. How effective was the method in covering the required domain areas or constructs (eg, usability, safety, workload, situational awareness, and decision-making)?
5. How effective was the method in generating recommendations that were useful and easy to implement?
6. How effective was the method in identifying micro, meso, and macro (systems) level considerations that have safety implications?
7. How effective was the method in understanding the dynamic nature of the system?
8. What was the impact on end user satisfaction, adoption, workflows, and workloads?

Qualitative evaluation, for example, via focus groups or surveys

Reflect on the overall HF approach including methods applied, considering the following:

1. What was the perceived overall effectiveness and impact with respect to safety and usability? How could these be improved?
2. What was the perceived overall effectiveness and impact on end user workflows and workloads? How can these be improved?

Funding

SA was supported by an Australian Government Research Training Program (RTP) scholarship.

Conflicts of Interest

None declared.

References

1. 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 7;160(1):48-54. [doi: 10.7326/M13-1531] [Medline: 24573664]
2. Agius S, Cassar V, Bezzina F, Topham L. Leveraging digital technologies to enhance patient safety. Health Technol 2025 Nov;15(6):1053-1063. [doi: 10.1007/s12553-025-01001-6]
3. Insani WN, Zakiyah N, Puspitasari IM, et al. Digital health technology interventions for improving medication safety: systematic review of economic evaluations. J Med Internet Res 2025 Feb 5;27:e65546. [doi: 10.2196/65546] [Medline: 39909404]
4. Adams KT, Pruitt Z, Kazi S, et al. Identifying health information technology usability issues contributing to medication errors across medication process stages. J Patient Saf 2021 Dec 1;17(8):e988-e994. [doi: 10.1097/PTS.0000000000000868] [Medline: 34009868]
5. ISO 9241-210 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 2025-05-25]
6. ISO 9241-143:2012(e) ergonomics of human-system interaction part 143: forms. International Organization for Standardization. 2012. URL: https://www.iso.org/standard/53590.html [accessed 2025-05-25]
7. Technical basis for user interface design of health IT. National Institute of Standards and Technology (NIST) 15-996. 2015. URL: https://www.nist.gov/publications/technical-basis-user-interface-design-health-it [accessed 2025-05-25]
8. Awad S, Amon K, Baillie A, Loveday T, Baysari MT. Human factors and safety analysis methods used in the design and redesign of electronic medication management systems: a systematic review. Int J Med Inform 2023 Apr;172:105017. [doi: 10.1016/j.ijmedinf.2023.105017] [Medline: 36809716]
9. Brown CL, Mulcaster HL, Triffitt KL, 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 Mar 1;24(2):432-440. [doi: 10.1093/jamia/ocw119] [Medline: 27582471]
10. Carayon P, Du S, Brown R, Cartmill R, Johnson M, Wetterneck TB. EHR-related medication errors in two ICUs. J Healthc Risk Manag 2017 Jan;36(3):6-15. [doi: 10.1002/jhrm.21259] [Medline: 28099789]
11. Ratwani RM, Savage E, Will A, 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]
12. Kushniruk A, Nohr C, Borycki E. Human factors for more usable and safer health information technology: where are we now and where do we go from here? Yearb Med Inform 2016 Nov 10;2016(1):120-125. [doi: 10.15265/IY-2016-024] [Medline: 27830239]
13. Cahill M, Cleary BJ, Cullinan S. The influence of electronic health record design on usability and medication safety: systematic review. BMC Health Serv Res 2025 Jan 6;25(1):31. [doi: 10.1186/s12913-024-12060-2] [Medline: 39762805]
14. Guidelines. National Health and Medical Research Council (NHMRC). URL: https://www.nhmrc.gov.au/guidelines [accessed 2025-06-15]
15. Dul J, Bruder R, Buckle P, et al. A strategy for human factors/ergonomics: developing the discipline and profession. Ergonomics 2012;55(4):377-395. [doi: 10.1080/00140139.2012.661087] [Medline: 22332611]
16. Ahlstrom V. Human factors design standard (DOT/FAA/HF-STD-001B). US Department of Transportation Federal Aviation Administration. 2016. URL: https://hf.tc.faa.gov/publications/2016-12-human-factors-design-standard/full_text.pdf [accessed 2025-05-25]
17. King BJ, Read GJM, Hulme A, et al. Evaluating the use of systems thinking methods in healthcare: a RE-AIM analysis of AcciMap and Net-HARMS. Ergonomics 2025 Sep;68(9):1542-1560. [doi: 10.1080/00140139.2024.2423170] [Medline: 39552189]
18. Waterson P, Robertson MM, Cooke NJ, Militello L, Roth E, Stanton NA. Defining the methodological challenges and opportunities for an effective science of sociotechnical systems and safety. Ergonomics 2015;58(4):565-599. [doi: 10.1080/00140139.2015.1015622] [Medline: 25832121]
19. Leveson N. A new accident model for engineering safer systems. Saf Sci 2004 Apr;42(4):237-270. [doi: 10.1016/S0925-7535(03)00047-X]
20. Salmon PM, King B, Hulme A, et al. Toward the translation of systems thinking methods in patient safety practice: assessing the validity of Net-HARMS and AcciMap. Saf Sci 2023 Mar;159:106003. [doi: 10.1016/j.ssci.2022.106003]

21. Carayon P, Salwei ME. Moving toward a sociotechnical systems approach to continuous health information technology design: the path forward for improving electronic health record usability and reducing clinician burnout. J Am Med Inform Assoc 2021 Apr 23;28(5):1026-1028. [doi: 10.1093/jamia/ocab002] [Medline: 33537756]
22. Hulme A, Stanton NA, Walker GH, Waterson P, Salmon PM. Are accident analysis methods fit for purpose? Testing the criterion-referenced concurrent validity of AcciMap, STAMP-CAST and AcciNet. Saf Sci 2021 Dec;144:105454. [doi: 10.1016/j.ssci.2021.105454]
23. Hulme A, Stanton NA, Walker GH, Waterson P, Salmon PM. Testing the reliability and validity of risk assessment methods in human factors and ergonomics. Ergonomics 2022 Mar;65(3):407-428. [doi: 10.1080/00140139.2021.1962969] [Medline: 34328389]
24. Stanton NA. Human Factors Methods: CRC Press.
25. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020;15(5):e0232076. [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]
26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234] [Medline: 34394658]
28. Cook RI. How complex systems fail. : University of Chicago; 2002 URL: https://how.complexsystems.fail/ [accessed 2025-05-25]
29. Dekker S. Drift into Failure: From Hunting Broken Components to Understanding Complex Systems: Ashgate Publishing Ltd; 2011.
30. Williams K. Safety, systems, complexity, and resilience: what makes organizations safe? In: Haxby E, Hunter D, Jaggar S, editors. An Introduction to Clinical Governance and Patient Safety: Oxford University Press; 2010.
31. Ibrahim Shire M, Jun GT, Robinson S. The application of system dynamics modelling to system safety improvement: present use and future potential. Saf Sci 2018 Jul;106:104-120. [doi: 10.1016/j.ssci.2018.03.010]
32. Qureshi ZH, Campbell A. Systemic safety and accident modelling of complex socio‐technical systems. INCOSE Int Symp 2009 Jul;19(1):21-35. [doi: 10.1002/j.2334-5837.2009.tb00935.x]
33. Zarei E, Khan F, Abbassi R. A dynamic human-factor risk model to analyze safety in sociotechnical systems. Process Saf Environ Prot 2022 Aug;164:479-498. [doi: 10.1016/j.psep.2022.06.040]
34. Resilience assessment based on models of functional resonance. Linköping University. 2008. URL: https://www.resilience-engineering-association.org/download/resources/symposium/symposium-2006/Woltjer_text.pdf [accessed 2025-05-25]
35. Developing and using program logic: a guide. NSW Health. 2023. URL: https://www.health.nsw.gov.au/research/Publications/developing-program-logic.pdf [accessed 2025-05-25]
36. Model for improvement: establishing measures. Institute for healthcare improvement. 2025 Feb 23. URL: https://www.ihi.org/learn/courses/open-school/catalog/qi-102 [accessed 2025-05-25]

Abbreviations

HF: human factors
HIT: health IT

Edited by A Choudhury; submitted 02.Mar.2025; peer-reviewed by J Poots, J Haun; revised version received 11.Dec.2025; accepted 26.Dec.2025; published 03.Feb.2026.

Please cite as:
Awad S, Begg R, Loveday T, Baillie A, Baysari M
Expert Views on Criteria for Evaluation of Human Factors Methods: Qualitative Interview Study
JMIR Hum Factors 2026;13:e73324
URL: https://humanfactors.jmir.org/2026/1/e73324
doi:10.2196/73324

Using the Theoretical Domains Framework to Identify Barriers to and Enablers of Patient Telemedicine Services Use in China: Qualitative Study

Ke Liu[1,2*], MPA; Yuting Yang[3*], MS; Zixuan Song[1,2], MS; Huixian Li[1,2], MS; Yanli Lyu[1,2], MS; Ke Zhang[2,4], MD; Xinxia Wu[1,2], MM; Zheng Hou[2,5], MD; Yipei Wang[3], MBA, PhD

[1]Department of Medical Affairs, Peking University Third Hospital, Beijing, China
[2]Office of Internet Hospital, Peking University Third Hospital, Beijing, China
[3]Institute of Hospital Management, Peking University Third Hospital, North Huayuan Road 49, Beijing, China
[4]Department of Otolaryngology, Peking University Third Hospital, Beijing, China
[5]Department of Gynecology and Obstetrics, Peking University Third Hospital, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Yipei Wang, MBA, PhD
Institute of Hospital Management, Peking University Third Hospital, North Huayuan Road 49, Beijing, China

Abstract

Background: Telemedicine has rapidly expanded worldwide due to its convenience and accessibility. In China, an increasing number of hospitals have begun offering telemedicine services; however, patient utilization remains relatively low. Limited research has examined patients' behaviors during the process of adopting telemedicine services.

Objective: This study aimed to identify barriers to and enablers for patients using telemedicine services and to formulate implementation strategies.

Methods: We conducted semistructured qualitative interviews based on the Theoretical Domains Framework (TDF) to identify barriers and enablers to telemedicine utilization. Twenty-one patients who had used Peking University Third Hospital's telemedicine services were included in the interviews. Data were analyzed using NVivo 12.0 with deductive thematic analysis guided by the TDF. Moreover, a group of experts was assembled to devise potential intervention strategies.

Results: A total of 28 themes were identified, including 14 barriers and 14 enablers across 5 of the 14 TDF domains. The most frequently reported barriers were operational challenges, prolonged waiting periods from asynchronous communication, and doubts about therapeutic efficacy, whereas the most frequently mentioned enablers were the convenience of telemedicine, time conservation, and support from hospitals. On the basis of these factors, we devised 6 intervention strategies.

Conclusions: This study demonstrated that patients' utilization of telemedicine services was affected by several barriers and enablers, including system architecture and design, patient interactions using telemedicine, and external assistance. To enhance the utilization, these factors must be meticulously considered. This study also suggests strategies to enhance the utilization of telemedicine.

(JMIR Hum Factors 2026;13:e78457) doi:10.2196/78457

KEYWORDS

telemedicine services; Theoretical Domains Framework; barriers; enablers; patients; qualitative study

Introduction

With the rapid advancement of information technology, telemedicine services have expanded globally [1,2]. In China, the provision of telemedicine services by medical institutions has increased significantly since the COVID-19 pandemic [3]. Telemedicine services have expanded patients' health care options [4]. Telemedicine is thought to enhance the efficiency of health care services [5,6], decrease medical costs [7,8], improve patient satisfaction [9], and facilitate the equitable distribution of medical resources worldwide [10,11]. Despite the growing number of institutions providing telemedicine, patient utilization remains low. As of 2023, more than 3000 institutions had provided telemedicine services, delivering nearly 100 million virtual visits. Despite the seemingly large service volume, telemedicine accounts for only 1% of all visits [12], a proportion further validated by patient interviews [13]. Conversely, almost 20% of medical consultations in the United States were performed via telemedicine in 2020 [14]. The establishment of telemedicine services without corresponding utilization has resulted in a significant waste of financial and human resources.

To enhance telemedicine utilization, it is crucial to identify and understand the factors influencing patient engagement. Previous studies have identified demographic factors, such as age [15], co-residence with children [16], and education [17], but these factors are difficult to change. Consequently, researchers turned to behavioral frameworks, such as the unified theory of acceptance and use of technology (UTAUT) [18], the technology acceptance model (TAM) [19], and the theory of planned behavior [20]. These frameworks have helped identify factors, such as perceived usefulness [21], ease of use [22], and social influence [23], as critical precursors to user intention. However, these models primarily focus on reflective processes, namely, how cognitive evaluations lead to the intention to use. Therefore, these classical models offer less granularity regarding the contextual barriers, habitual behaviors, and skill gaps that prevent intention from translating into actual usage in real-world health care environments.

The Theoretical Domains Framework (TDF) addresses this gap not by discarding prior theories but by synthesizing them into a more granular, intervention-oriented structure rooted in the capability, opportunity, and motivation system [24]. Thus, adopting the TDF allows for a true integration of classical constructs: for instance, the "perceived ease of use" in TAM is no longer just a perception but is further defined in operational context as "capability" (ie, the knowledge and skills domains); "social influence" in the UTAUT and the theory of planned behavior is recontextualized within "opportunity" (ie, the social influences domain); and "perceived usefulness" aligns well with "motivation" (ie, the beliefs about capabilities and beliefs about consequences domains) [25].

Therefore, the TDF effectively extends the theoretical boundary beyond the cognitive focus of TAM and UTAUT. By incorporating less-explored domains, such as environmental context and resources, the TDF captures the contextual drivers of behaviors that traditional models often overlook [26,27]. Its comprehensiveness enables our research to move beyond a simple prediction of adoption intention. Instead, it provides a diagnostic structure to identify relevant barriers and enablers affecting telemedicine usage [28,29]. Adopting the TDF thus situates our work within the existing literature, bridging the well-established cognitive constructs of TAM and UTAUT with the implementation-focused depth required for designing interventions aimed at improving patient utilization of telemedicine services in China.

Building on this theoretical foundation, this study targets patients who have used telemedicine services in China, aiming to examine their behaviors and psychological aspects during the telemedicine process. It used semistructured qualitative interviews based on the TDF to ascertain barriers and enablers to telemedicine adoption. Our research may offer actionable evidence to support the refinement of telemedicine platforms and the development of patient-centered service models in future practice.

Methods

Study Design

We conducted the survey in Beijing, a leading Chinese metropolis in the development of telemedicine. This study focused on tertiary public hospitals, which deliver the majority of telemedicine services in China and are highly trusted and sought after by patients [30,31]. Peking University Third Hospital, a tertiary public hospital, serves as the governing institution of the Beijing Telemedicine Quality Control Center [32]. It ranks among the top 10 hospitals in Beijing in terms of telemedicine user volume, with patients coming from dozens of provinces nationwide over recent years. The hospital provides both synchronous and asynchronous telemedicine consultations, covering the 2 predominant service models in China. Its extensive online services span the entire care process, including prediagnosis, diagnosis, and postdiagnosis.

Participants Recruitment

We used convenience sampling to select patients engaged in real-time video and asynchronous graphic consultations at Peking University Third Hospital. We accounted for patients' varied geographic locations, disease classifications, and socioeconomic statuses to ensure that the sample accurately represented the telemedicine-using patient community. The inclusion criteria were as follows: (1) at least one prior use of telemedicine at Peking University Third Hospital within the past 6 months, (2) willingness to participate in the interview, and (3) adequate communication skills to engage in the interview. Previous literature suggests that thematic saturation in qualitative interview studies is generally achieved before reaching 21 interviews [33]. Accordingly, the intended sample size was 20 to 22 participants in the preliminary phase. When the data collection process yielded no new information about the enablers and barriers influencing access to telemedicine services, data saturation was reached [34].

Interview Guide Development

We applied the TDF throughout the interview design, data collection, and analysis. The TDF guided the development of the interview guide, which included 1 to 3 questions for each of the 14 domains (eg, goals and intentions) [35]. For example, one of the questions was, "What do you know about telemedicine services?" Investigators with expertise in implementation science provided guidance to develop the interview guide. Moreover, we consulted with experts in hospital management, health policy management, and physicians to refine the outline. We conducted 2 preexperiments after the outline was initially finalized. The interview outline was refined slightly after the pretests, and the samples from the pretests were not included in the study due to slight changes. Table 1 presents the TDF domains and corresponding interview questions.

Table . Interview questions aligned with the Theoretical Domains Framework (TDF) domains.

TDF domains	Definition [26]	Definition in this study	Interview questions
Goals	Mental representations of outcomes or end states that an individual wants to achieve	Goals that patients hope to resolve using telemedicine	• What problems can telemedicine help you solve?
Intention	A conscious decision to perform a behavior or resolve to act in a certain way	Patients' willingness or plan to use telemedicine	• How did you know about telemedicine, and why did you try it? • How do you feel about it compared to the offline service?
Knowledge	An awareness of the existence of something	The patient's understanding of telemedicine	• What do you know about telemedicine services?
Skills	An ability or proficiency acquired through practice	The patient's practical ability to operate telemedicine independently	• Are you familiar with using telemedicine services? • Can you independently use telemedicine for online consultations?
Beliefs about capabilities	Acceptance of the truth, reality, or validity about an ability, talent, or facility that a person can put to constructive use	Patients' confidence in their ability to use telemedicine successfully	• To what extent do you feel capable of using telemedicine? • When you use telemedicine, is there someone to help you? If so, who and how?
Optimism	The confidence that things will happen for the best or that desired goals will be attained	Patients' positive expectations about the benefits of telemedicine	• Do you think telemedicine can benefit patients? • Is it beneficial to you? Please provide some examples.
Memory, attention, and decision processes	The ability to retain information, focus selectively on aspects of the environment, and choose between 2 or more alternatives	Cognitive aspects influencing telemedicine patients' memory, concentration, and decision-making	• Are there aspects that complicate the utilization of telemedicine services? Have you faced any particular challenges?
Emotions	A complex reaction pattern, involving experiential, behavioral, and physiological elements, by which the individual attempts to deal with a personally significant matter or event	Patients' emotional experiences when engaging with telemedicine	• How do you feel when you use telemedicine services? • How do you feel about telemedicine services compared to offline service?
Beliefs about consequences	Acceptance of the truth, reality, or validity of outcomes of a behavior in a given situation	Patients' perceptions of the benefits and risks of telemedicine	• What are the enablers of telemedicine services? • What are the barriers to telemedicine services?
Reinforcement	Increasing the probability of a response by arranging a dependent relationship, or contingency, between the response and a given stimulus	Incentives or feedback that motivate patients to use telemedicine	• What incentives can motivate you to use telemedicine services?
Social or professional role and identity	A coherent set of behaviors and displayed personal qualities of an individual in a social or work setting	Patients' perception of their role in health management and its alignment with telemedicine	• Have people around you (family members, friends, etc) used telemedicine services? • What attitudes do they hold about telemedicine services? • Does their perspective affect your decision regarding the visit?
Social influences	Those interpersonal processes that can cause individuals to change their thoughts, feelings, or behaviors	Impact of family, peers, and social norms on telemedicine use	• Do social environment and policies affect your use of telemedicine services? How?

TDF domains	Definition [26]	Definition in this study	Interview questions
Environmental context and resources	Any circumstance of a person’s situation or environment that discourages or encourages the development of skills and abilities, independence, social competence, and adaptive behavior	External conditions and resources that facilitate or hinder telemedicine use	• Do you believe that the social environment and policies will influence your utilization of telemedicine services? What is its impact? • Does the availability of resources and support affect your use of telemedicine? How?
Behavioral regulation	Anything aimed at managing or changing objectively observed or measured actions	Patients’ strategies to ensure effective use of telemedicine	• What changes are necessary for telemedicine services to succeed? • Do you have any other suggestions to enhance telemedicine usage?

Interview and Data Collection

The formal interview was performed via telephone and approved in January 2024. We secured the participants’ informed consent and subsequently audio recorded the interviews using the software’s exclusive features. Each interview lasted approximately 10 to 20 minutes. The interviews were conducted in Mandarin Chinese by the authors KL and YY, both of whom had professional training in qualitative interviewing and extensive experience in qualitative research (n=2, one male and one female, health management practitioners). Before each interview, the interviewer provided a brief self-introduction and overview of the interview content to establish rapport with the participant. Only 2 of the interviewers were present during the interview to create a conducive environment for communication. The recordings were transcribed into textual materials within 48 hours after the interviews, and the researcher recorded and sorted the data to ensure consistency. After each interview, participants were sent a prearranged text message, and their data were collected.

Data Coding and Analysis

Interview transcripts were transcribed, organized, and imported into the software NVivo (version 12.0; QSR International Pty Ltd) for coding. We conducted a deductive thematic analysis guided by the TDF and followed 3 steps [36,37]. First, the 14 TDF domains served as an a priori coding framework. Second, the transcripts were reviewed in detail, with relevant content identified and coded into themes. Third, these themes were categorized as either barriers or enablers within the corresponding TDF domains. When a theme was relevant to more than 1 TDF domain, it was assigned to the domain judged to be more closely aligned. Two authors (KL and YY) independently coded the transcripts. Coding discrepancies were resolved through discussion with a third author (YW) until consensus was reached. The coded κ for each level was greater than 0.90. To prepare the results, we tabulated the themes within each TDF domain and tabulated more prominent themes into barriers and enablers separately. A detailed description of the coding process, including example excerpts and coding framework, was provided in Multimedia Appendix 1.

Strategy Development

Following encoding and analysis, we assembled a multidisciplinary panel of hospital experts, including specialists in information technology, clinical practice, and hospital administration leadership. This collaborative endeavor integrated technical, clinical, and operational viewpoints via monthly interdepartmental meetings and methodically structured brainstorming sessions employing implementation science frameworks. Using the TDF model, the panel systematically formulated various intervention options aimed at addressing the identified barriers and enablers.

Ethical Considerations

This study was approved by the Medical Science Research Ethics Committee of Peking University Third Hospital (IRB00006761-M2024029). Participants provided informed consent. We stored the survey data, which were strictly managed to remain confidential. We offered a compensation of 100 yuan (US $13.8) as a token of appreciation to all participants for their involvement in the study.

This study followed established reporting standards for qualitative research and was reported in accordance with the COREQ checklist (Checklist 1).

Results

Participants and Institutions

All interviews were conducted from January to March 2024, covering a duration of 3 months. Twenty-five telephone interviews were conducted, of which 21 were considered legitimate for analysis. Of the 21 interviewees, 13 (62%) were women, and 19 (91%) held a bachelor’s degree or higher. The majority (18/21, 85%) of the participants had full-time jobs. The mean interview duration was 13.7 (SD 5.7) minutes. All participants had used telemedicine services at least once, with a mean of 3.46 (SD 2.8) visits as specified in Table 2.

Table . Interview patient characteristics (N=21).

Characteristics	Participants, n (%)
Sex	
Male	8 (38)
Female	13 (62)
Age (y)	
<30	4 (19)
30 - 35	5 (24)
36 - 40	7 (33)
>40	5 (24)
Education	
Senior high school or below	2 (9)
College	10 (48)
Postgraduate degree or above	9 (43)
Employment status	
Full-time employment	18 (85)
Unemployed	1 (5)
Retired	1 (5)
Student	1 (5)
Home location	
Beijing	13 (62)
Outside of Beijing	8 (38)
Having a chronic disease	
Yes	3 (15)
No	18 (85)
Undergone surgery	
Yes	3 (15)
No	18 (85)
Medical insurance type	
Employee insurance	14 (67)
Publicly funded free medical care	6 (28)
Urban resident insurance	1 (5)
Number of uses	
1	8 (38)
2 - 5	8 (38)
6 - 10	5 (24)

Barriers to and Enablers of Patient Use of Telemedicine Services

A total of 190 principal assertions were documented, covering 5 domains and 28 themes, including 14 barriers and 14 enablers. Tables 3 and 4 display the themes within each domain and their frequencies.

Table . Barriers to use of telemedicine services (n=71).

TDF[a] domains and themes	Frequency, n (%)
Environmental context and resources (n=33)	
The interaction of the app is poor	12 (17)
The functions of the telemedicine services are limited	8 (11)
The process of telemedicine services and treatment is not smooth enough	5 (7)
The operation is not intelligent enough and is rather cumbersome	4 (6)
The propagandizing is insufficient	3 (4)
The environment is too noisy	1 (1)
Memory, attention, and decisions process (n=23)	
Patients have to wait passively for a long time by asynchronous communication	10 (14)
The operation is too complicated for the patients	10 (14)
It is hard to find the functions that suit oneself	3 (4)
Beliefs about consequences (n=8)	
Doubts and distrust about the effectiveness of telemedicine services	7 (10)
The risk of privacy leakage and security issues	1 (1)
Skills (n=6)	
Mobile phone operation skill is poor	4 (6)
It is difficult to collect information online	2 (3)
Beliefs about capabilities (n=1)	
Patients have no confidence in telemedicine services	1 (1)

[a] TDF: Theoretical Domains Framework.

Table . Enablers to use of telemedicine services (n=119).

TDF[a] domains and themes	Frequency, n (%)
Beliefs about consequences (n=67)	
The convenience of telemedicine services	36 (30)
Telemedicine services can save time	16 (13)
Telemedicine services can enhance appointment accessibility	10 (8)
Telemedicine services can reduce transportation costs	5 (4)
Environmental context and resources (n=37)	
The system is easy to operate	10 (8)
Hospitals or app can provide guidance	8 (7)
Stable and fast network	8 (7)
Provide prompt consultation services	4 (3)
There is a quiet and suitable environment	4 (3)
Has the function of prescribing medicine	3 (3)
Beliefs about capabilities (n=8)	
The patient has self-efficacy in digital health	5 (4)
The patient has successful experience in online diagnosis and treatment	3 (3)
Skills (n=7)	
The patient's mobile phone operation skills are relatively good	5 (4)
The patient has the ability to collect information	2 (2)

[a] TDF: Theoretical Domains Framework.

The barriers to theoretical domains were referenced a total of 71 times. Specifically, “environmental context and resources” was referenced 33 times; “memory, attention, and decision processes” was referenced 23 times; “beliefs about consequences” was referenced 8 times; “skills” was referenced 6 times; and “beliefs about capabilities” was referenced once. In contrast, the enablers of theoretical domains were referenced a total of 119 times. “Beliefs about consequences” was referenced most frequently, occurring 71 times”; “environmental context and resources” was referenced 37 times; “beliefs about capabilities” was referenced 8 times; and “skills” was referenced 7 times.

Barriers

Environmental Context and Resources

“Environmental context and resources” was the primary barrier. It primarily emphasized user-environment interaction and the functional categories of telemedicine services. On the one hand, user-environment interaction explicitly examined how the design attributes of app influence patient engagement. It encompassed interface usability and system intelligence. On the other hand, the limited variety of functions created barriers. If patients' medical needs were not fully met, they did not use telemedicine services.

> *The software design seems flawed, insufficient prompts and guidance mechanisms make navigation unnecessarily challenging.* [LQ, male, 2-time user]

> *The functions of online consultation are too limited. It would be great if they could be more diverse to meet more demands.* [KJ, female, 8-time user]

Memory, Attention, and Decision Process

The patients' memory, attention, and decision-making processes might cause barriers, including asynchronous communication and the complexity of the process. First, barriers arose from asynchronous communication between patients and physicians. Following patients' inquiries, they frequently endured prolonged delays before receiving a response from the physician, which significantly impacted patients' readiness to use telemedicine. Second, barriers emerged from the complexity of the process. The complexity of the process was referenced 8 times in the interview, exclusively by first-time users. Patients' attention was limited, but too many steps and jumps caused patients to resist the use of telemedicine services.

> *Sometimes we want to contact the physician proactively and leave a message there, but the physician can't see it. We can only wait for the physician to contact us. This is too passive.* [LJT, female, 2-time user]

> *I searched endlessly for the payment page but couldn't find it. The design isn't user-friendly.* [YY, female, 1-time user]

Beliefs About Consequences

“Beliefs about consequences” manifested as barriers from two primary concerns: skepticism about treatment effectiveness and fears of privacy infringements. Primarily, owing to inadequate communication during telemedicine consultations, patients doubted the efficacy of telemedicine services. In addition, telemedicine necessitated the accumulation of extensive patient data, which led to concerns regarding the potential breach of personal privacy.

> *I believe that online medical consultations are not as reliable as in-person visits. In a physical setting, physicians can conduct thorough examinations, whereas online consultations rely solely on verbal descriptions.* [WWH, male, 1-time user]
>
> *I am genuinely concerned about the potential for my personal information or medical data to be leaked. Given the prevalence of telecom fraud, I prefer to go to hospital.* [YYX, male, 3-time user]

Skills

"Skills" emerged as a critical barrier domain, primarily manifested through deficits in digital health literacy and device operation proficiency. During the interviews, numerous patients indicated that telemedicine presented specific challenges for the older adults.

> *At our age, we're barely comfortable with basic smartphone functions like WeChat. Navigating these sophisticated medical platforms feels overwhelmingly complex.* [ST, female, 3-time user]

Enablers

Beliefs About Consequences

"Beliefs about consequences" was repeatedly highlighted as an enabler, reflecting patients' recognition of the benefits and potential outcomes of telemedicine. First, nearly all polled patients indicated that convenience facilitated their utilization of telemedicine services, with several patients reiterating this point consistently. Second, time conservation emerged as a key enabler, as telemedicine significantly reduced time expenditure, particularly for younger patients. Third, improved appointment accessibility was identified as a crucial enabler. Patients were able to schedule follow-up tests and obtain medications via telemedicine services without concerns regarding registration.

> *After we use it, we think it is really convenient, so we don't have to go to the hospital anymore, and it is not easy for us to go there.* [YY, female, 1-time user]
>
> *I think using this [telemedicine services] can communicate with the physician, the effect is similar, and it is much easier than offline, no need to ask for leave or go to the hospital.* [SST, female, 1-time user]
>
> *I've found departments are readily available on the telemedicine services platform. That's really helpful.* [TXM, female, 7-time user]

Environmental Context and Resources

"Environmental context and resources" was identified as an enabler influencing patients' use of telemedicine services. First, app usability emerged as an enabler, as patients typically perceived that simple operation enhanced their engagement with the telemedicine services. Second, support from the app or hospital proved essential, helping patients easily locate access to telemedicine services and substantially improving utilization. Finally, the comprehensiveness of platform functionalities was recognized as a crucial enabler, as an extensive range of telemedicine services facilitated patients' adoption and use.

> *Telemedicine services are straightforward. I like it.* [WZ, male, 5-time user]
>
> *The hotline service is particularly helpful. Whenever I encounter difficulties, their staff guides me through the process step-by-step.* [LJT, female, 2-time user]
>
> *We hope they'll keep expanding features so we don't have to make endless trips to the hospital anymore.* [PY, male, 2-time user]

Beliefs About Capabilities

"Beliefs about capabilities" evolved as enablers, reflected in two aspects: self-efficacy and prior successful experience. On the one hand, self-efficacy in digital health was largely associated with usage frequency, as expressed by confident users. On the other hand, patients' previous successful experiences influenced their utilization behaviors, as they drew upon earlier eHealth encounters.

> *I'm generally good at figuring out tech stuff. People around me always ask me for advice when they need online medical help.* [YY, female, 1-time user]
>
> *Having used other hospitals' internet services, I assume this APP should work fairly similarly.* [KJ, female, 8-time user]

Skills

"Skills" emerged as a crucial enabler in the use of telemedicine services. In particular, expertise in using mobile phones and information retrieval was frequently mentioned. Individuals with advanced mobile device skills were more inclined to use telemedicine services. In addition, proficiency in information-seeking behaviors was recognized as an enabler.

> *I'm quite comfortable using smartphones daily, so navigating these platforms seems easy.* [LLZ, male, 1-time user]
>
> *I independently researched all relevant information about telemedicine services. Honestly, I'm pretty good at finding what I need.* [YY, female, 1-time user]

Interventions to Improve Patient Utilization of Telemedicine Services

On the basis of these findings, we designed intervention measures addressing barriers and enablers through expert group discussions (Table 5). We assessed the importance of every domain by determining the frequency of citations throughout all transcripts. The domains "beliefs about consequences" and "environmental context and resources" were the most frequently cited, with 75 and 70 mentions, respectively, accounting for 40% and 37% of the total citations.

Table . Evidence-based interventions to improve telemedicine services utilization (n=190).

TDF[a] domains	Frequency, n (%)	Intervention	Implementation details
Beliefs about consequences	75 (39)	Enhance support for first-time users	• Pilot postdischarge coaching in orthopedics • Train family physicians to assist virtual visits
Environmental context and resources	70 (37)	Simplify app functions	• Suspend development of complex features • Optimize core functionalities • Streamline the patient's usage process
Memory, attention, and decision process	23 (12)	Enhance APP guidance and 24/7 support hotline	• Add real-time operation prompts • Enable automatic step progression • Develop an AI[b]-powered patient assistant • Provide step-by-step guidance • Establish rapid complaint resolution • Implement AI-assisted feedback tracking
Skills	13 (7)	Multichannel publicity	• Distribute instructional videos via WeChat or social media • Install interactive kiosks in clinics • Create a physician's promotional webpage
Beliefs about capabilities	9 (5)	Age-friendly adaptation	• Launch large-text interface • Create elderly-specific tutorials • Integrate voice command features

[a] TDF: Theoretical Domains Framework.

[b] AI: artificial intelligence.

In the "beliefs about consequences" domain, convenience emerged as a key factor encouraging patients to use telemedicine. Enhancing the experiences of first-time users was considered essential. It was advisable to offer supplementary services for patients using telemedicine for the first time, thereby enhancing their experience of its convenience.

Furthermore, in the "environmental context and resources" domain, the complexity of using the patient application represented the most prominent barrier factor. It was suggested to simplify the patient interface and limit the proliferation of complex features, hoping to facilitate patients' focus on critical functions to meet the needs of the majority.

Discussion

Principal Findings

This study conducts an in-depth exploration of the relatively underresearched area of telemedicine among Chinese patients. It is the first instance of using the TDF model for conducting interview-based surveys with patients regarding telemedicine, thereby enhancing the scientific rigor of the research. This study found that telemedicine services adoption is jointly shaped by 14 barriers and 14 enablers, including 5 of the 14 TDF domains. This study has found more behavioral and psychological influencing factors in patients compared to other studies. Barriers arose from operational complexity, concerns about privacy and efficacy, asynchronous communication, and so on. Enablers include high convenience and accessibility, simplified system functionalities, and external support. On the basis of identified barriers and enablers, we developed targeted intervention measures and prioritized them based on their relative significance as expressed by the interviewees. In conclusion, this study advances a comprehensive understanding of the determinants influencing this adoption and establishes a theoretical foundation for developing targeted interventions in telemedicine services.

Positive Feedback Mechanisms

Our exploratory analysis of usage frequency indicates that barriers and enablers change alongside patients' cumulative experiences with telemedicine services. First-time or infrequent users were more inclined to highlight challenges associated with accessing the platform and operating system, indicating the significant learning curve of initial usage. In contrast, frequent users recognized constraints in the scope and profundity of accessible system functionalities while concurrently emphasizing convenience and efficiency as primary enablers. This transition suggests a possible positive feedback loop: as patients become accustomed to the system, operational obstacles lessen, and the perceived advantages of telemedicine become more prominent, hence encouraging further utilization.

This corresponds with previous research [13,38]. Similarly, in the TAM, it is believed that "perceived usefulness" influences behavioral intention [39]. Our research delineates a self-perpetuating feedback loop: utilization of telemedicine services leads to increased satisfaction and reinforced intention to use. This phenomenon highlights the importance of the first usage experience in telemedicine service adoption. Similarly, the anchoring effect in initial interactions profoundly influences long-term commitment. Facilitating smooth navigation and practical convenience during initial interactions might create positive behavioral anchors, promoting regular use of online telemedicine services. Previous research indicates that Michigan Medicine (Ann Arbor) has established a system that incorporates template instructions and anticipatory advice, which can enhance the patient utilization process [40]. Consequently, we advocate for the deployment of multimodal guiding systems, including artificial intelligence–driven triage and a human-assisted hotline.

Duality of Barriers and Enablers

Our study reveals that the utilization of telemedicine services is directionally modulated by coexisting enablers and barriers within the same theoretical domains. The 2 domains, "environmental context and resource" and "the beliefs about consequences," involve many barriers and enablers. Within the identical structural domain, both barriers and enablers arose, and the correlation between the two was notably strong.

The initial point is the discrepancy between functional complexity and operational simplicity. Within the "environmental context and resource" domain, patients simultaneously demand operational simplicity and comprehensive functionalities. To address patients' requirements of operational simplicity, we must reduce the challenges to assure accessibility. This aligns with the findings of our and other research from the viewpoint of physicians [25,41,42]. The complexity of system operation significantly affects both physicians and patients. However, excessive feature expansion paradoxically increases interface complexity, hindering core functionalities. This phenomenon aligns with intrinsic cognitive load theory [43]. In the development of telemedicine services, it is imperative to address this equilibrium. This is a critical aspect that we meticulously consider while devising potential intervention strategies.

The second aspect is the trade-off between convenience and privacy. The "beliefs about consequences" domain encompasses opposing factors: perceived convenience and privacy concerns. In the study, the convenience of telemedicine services is the most frequently mentioned enabler by patients. However, this depends on patients supplying comprehensive personal information, medical history, and privacy. This also results in an elevated danger of personal privacy breaches. Although the issue of privacy breaches is only mentioned once in this study, it is frequently cited as a barrier in many studies on the use of telemedicine services [44]. Patients tend to resist using telemedicine services due to the risk of privacy breaches [45]. This also results in a simultaneous presence of dependence on convenience and sensitivity to contradictions, leading to conflicting and hesitant feelings in patients during the usage process [46].

The Mismatch Between the Demand and Usage Capacity

In skills, digital literacy is a significant influencing element. Numerous interviewers highlighted the concern of digital literacy for older adults, despite the interviewers in the study being youthful. Older patients, despite increasing demand resulting from mobility restrictions and multimorbidity, face barriers due to limited digital literacy and adaptability [47,48]. In contrast, younger groups with technical skills demonstrate reduced demand. This generational mismatch hinders the use of telemedicine services by patients.

In establishing intervention strategies for telemedicine services, it is essential to address the requirements and features of both generations, rather than focusing on only one group. The alteration of text size was frequently referenced during the interview. Consequently, we devised an app modification strategy for the older people, offering extensive assistance while preserving the core functionality. We have also developed the "family affection account" feature, which enables children to assist the older adults in completing telemedicine-related processes.

Legal and Security Concerns

In research primarily focusing on Western populations, privacy concerns are often cited as a major barrier to adoption [41,49-51]. Yet, this barrier was not prominent in our study, with only one participant explicitly raising privacy-related concerns. We speculate on the possible reasons from both macro levels and micro levels.

At the macro level, the social discourse and privacy awareness in China and Western countries differ markedly. In Europe and the United States, privacy protection is widely discussed [52], whereas in China, discussions regarding medical data privacy are limited, and most patients have low awareness, leading to infrequent expression of privacy concerns [52,53].

At the micro level, Chinese patients typically exhibit a high degree of trust in health care institutions, largely due to their public and government-regulated nature [54], which may further reduce concerns about potential data misuse [55].

Limitations of the Study

This study has 3 main limitations. First, the sample consisted only of individuals who had already used telemedicine and were recruited from a single tertiary public hospital in Beijing, which may limit the generalizability of the findings. Future studies should consider including more diverse populations. Second, using a deductive approach to map participants' statements to TDF domains may result in inconsistencies across coders. To mitigate this inherent limitation of deductive analysis, we used double independent coding, with any discrepancies resolved through adjudication by a third coder. Given the high consistency of participant narratives regarding barriers and enablers, we believe any residual impact is minimal. Third, as this study represents the first paper in a planned series, its primary goal was to identify barriers to and enablers of telemedicine adoption, thereby laying the foundation for intervention development. Thus, this study did not assess the effectiveness of proposed interventions. In future work, we will incorporate considerations of feasibility and other factors to prioritize specific interventions and rigorously evaluate their effectiveness.

Conclusions

This study offered vital insights by identifying barriers to and enablers of patients using telemedicine services based on the TDF model. These findings not only genuinely represent patients' experiences but also reflect the behavioral and psychological factors in using telemedicine services. This study revealed barriers, including operational difficulties, extended waiting times, and doubts concerning therapeutic effectiveness. Simultaneously, the enablers included the convenience of telemedicine, time conservation, and support from hospitals. Considering these criteria, we developed intervention strategies to enhance patients' access to telemedicine. This research offers a solid foundation for developing more targeted intervention strategies and holds substantial practical implications for improving the telemedicine experience for patients.

Acknowledgments

The authors would like to appreciate the generous time that the interviewed participants gave during data collection. The authors sincerely thank the reviewers for their constructive comments and valuable suggestions, which greatly improved the quality of this manuscript.

Funding

This work was supported by the National Natural Science Foundation (72204005) and Capital's Funds for Health Improvement and Research under grant number 2024-2-40915.

Data Availability

The datasets used in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

KL and YY are co-first authors of this manuscript and contributed equally to this research. KL, YY, and YW planned and designed the study. KL and YY conducted the interview and collected and coded the data. ZS, HL, and YL assisted in organizing the interview records. KZ, XW, and ZH contributed the administrative and technical support of the study. KL, YY, and YW wrote the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The coding process and results (excerpt).
[DOCX File, 24 KB - humanfactors_v13i1e78457_app1.docx]

Checklist 1
COREQ checklist.
[DOCX File, 20 KB - humanfactors_v13i1e78457_app2.docx]

References

1. Klaassen B, van Beijnum BJF, Hermens HJ. Usability in telemedicine systems-a literature survey. Int J Med Inform 2016 Sep;93:57-69. [doi: 10.1016/j.ijmedinf.2016.06.004] [Medline: 27435948]
2. Zhang X, Lai KH, Guo X. Promoting China's mHealth market: a policy perspective. Health Policy Technol 2017 Dec;6(4):383-388. [doi: 10.1016/j.hlpt.2017.11.002]
3. Hong Z, Li N, Li D, et al. Telemedicine during the COVID-19 pandemic: experiences from Western China. J Med Internet Res 2020 May 8;22(5):e19577. [doi: 10.2196/19577] [Medline: 32349962]
4. Wang Y, Zhang P, Xing Y, et al. Telemedicine integrated care versus in-person care mode for patients with short stature: comprehensive comparison of a retrospective cohort study. J Med Internet Res 2024 Nov 19;26:e57814. [doi: 10.2196/57814] [Medline: 39560983]
5. Carrillo de Albornoz S, Sia KL, Harris A. The effectiveness of teleconsultations in primary care: systematic review. Fam Pract 2022 Jan 19;39(1):168-182. [doi: 10.1093/fampra/cmab077] [Medline: 34278421]

6. Peters GM, Kooij L, Lenferink A, van Harten WH, Doggen CJM. The effect of telehealth on hospital services use: systematic review and meta-analysis. J Med Internet Res 2021 Sep 1;23(9):e25195. [doi: 10.2196/25195] [Medline: 34468324]
7. Elbert NJ, van Os-Medendorp H, van Renselaar W, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014 Apr 16;16(4):e110. [doi: 10.2196/jmir.2790] [Medline: 24739471]
8. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020 Oct 19;22(10):e17298. [doi: 10.2196/17298] [Medline: 33074157]
9. Tan H, Zhang X, Yang Y. Satisfaction or gratitude? Exploring the disparate effects of physicians' knowledge sharing on patients' service evaluation in online medical consultations. Inf Syst J 2023 Sep;33(5):1186-1211. [doi: 10.1111/isj.12440]
10. Wang Z, Wang Y, Qiu Y, Jin C. Strategic resource allocation in dual-channel healthcare systems: the role of telemedicine in physician assignment. IEEE Trans Autom Sci Eng 2025;22:13569-13587. [doi: 10.1109/TASE.2025.3549822]
11. Caulfield BM, Donnelly SC. What is connected health and why will it change your practice? QJM 2013 Aug;106(8):703-707. [doi: 10.1093/qjmed/hct114] [Medline: 23676416]
12. Statistical bulletin on the development of China's health and wellness industry in 2023 [Report in Chinese]. : National Health Commission of the People's Republic of China; 2024 URL: https://www.yi123.cc/doc/7e50qoexjc [accessed 2026-01-09]
13. Wang S, Wang X, Zhou Y, Xu J. Utilization of, satisfaction toward, and challenges for Internet-based healthcare services provided by primary health institutions: evidence from China. Front Public Health 2023;10:1100634. [doi: 10.3389/fpubh.2022.1100634] [Medline: 36743153]
14. Weiner JP, Bandeian S, Hatef E, Lans D, Liu A, Lemke KW. In-person and telehealth ambulatory contacts and costs in a large US insured cohort before and during the COVID-19 pandemic. JAMA Netw Open 2021 Mar 1;4(3):e212618. [doi: 10.1001/jamanetworkopen.2021.2618] [Medline: 33755167]
15. Zhou M, Zhao L, Kong N, Campy KS, Qu S, Wang S. Factors influencing behavior intentions to telehealth by Chinese elderly: an extended TAM model. Int J Med Inform 2019 Jun;126:118-127. [doi: 10.1016/j.ijmedinf.2019.04.001] [Medline: 31029253]
16. de Veer AJE, Peeters JM, Brabers AEM, Schellevis FG, Rademakers JJDJM, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15:103. [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
17. AlBar AM, Hoque MR. Patient acceptance of e-Health services in Saudi Arabia: an integrative perspective. Telemed J E Health 2019 Sep;25(9):847-852. [doi: 10.1089/tmj.2018.0107] [Medline: 30452333]
18. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117. [doi: 10.2196/13117] [Medline: 31115340]
19. Lee AT, Ramasamy RK, Subbarao A. Understanding psychosocial barriers to healthcare technology adoption: a review of TAM technology acceptance model and unified theory of acceptance and use of technology and UTAUT frameworks. Healthcare (Basel) 2025 Jan 27;13(3):250. [doi: 10.3390/healthcare13030250] [Medline: 39942440]
20. Li D, Hu Y, Pfaff H, et al. Determinants of patients' intention to use the online inquiry services provided by internet hospitals: empirical evidence from China. J Med Internet Res 2020 Oct 29;22(10):e22716. [doi: 10.2196/22716] [Medline: 33006941]
21. Harris MA, Freeman KA, Duke DC. Seeing is believing: using Skype to improve diabetes outcomes in youth. Diabetes Care 2015 Aug;38(8):1427-1434. [doi: 10.2337/dc14-2469] [Medline: 26033508]
22. Green SM, Lockhart E, Marhefka SL. Advantages and disadvantages for receiving internet-based HIV/AIDS interventions at home or at community-based organizations. AIDS Care 2015;27(10):1304-1308. [doi: 10.1080/09540121.2015.1051503] [Medline: 26357907]
23. Li P, Luo Y, Yu X, et al. Patients' perceptions of barriers and facilitators to the adoption of e-Hospitals: cross-sectional study in Western China. J Med Internet Res 2020 Jun 11;22(6):e17221. [doi: 10.2196/17221] [Medline: 32525483]
24. Xu P, Jin Y, Guo P, et al. Barriers and enablers of pelvic floor rehabilitation behaviours in pregnant women with stress urinary incontinence: a qualitative analysis using the Theoretical Domains Framework. BMC Pregnancy Childbirth 2023 Apr 28;23(1):300. [doi: 10.1186/s12884-023-05633-2] [Medline: 37118702]
25. Wu X, Yang Y, Li Y, et al. Using theory-based frameworks to identify barriers and enablers of physicians' telemedicine adoption and develop intervention strategies in China: multicenter qualitative study. J Med Internet Res 2025 Sep 8;27:e73412. [doi: 10.2196/73412] [Medline: 40920450]
26. Cane J, O'Connor D, Michie S. Validation of the Theoretical Domains Framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37. [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
27. Shachak A, Kuziemsky C, Petersen C. Beyond TAM and UTAUT: future directions for HIT implementation research. J Biomed Inform 2019 Dec;100:103315. [doi: 10.1016/j.jbi.2019.103315] [Medline: 31629923]
28. Atkins L, Francis J, Islam R, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77. [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]

29. Yamada J, Kouri A, Simard SN, Segovia SA, Gupta S. Barriers and enablers to using a patient-facing electronic questionnaire: a qualitative Theoretical Domains Framework analysis. J Med Internet Res 2020 Oct 8;22(10):e19474. [doi: 10.2196/19474] [Medline: 33030437]
30. Han Y, Lie RK, Guo R. The internet hospital as a telehealth model in China: systematic search and content analysis. J Med Internet Res 2020 Jul 29;22(7):e17995. [doi: 10.2196/17995] [Medline: 32723721]
31. Xu D, Zhan J, Cheng T, Fu H, Yip W. Understanding online dual practice of public hospital doctors in China: a mixed-methods study. Health Policy Plan 2022 Apr 12;37(4):440-451. [doi: 10.1093/heapol/czac017] [Medline: 35266518]
32. Peking University Third Hospital has been approved as the chair unit of the Beijing internet-based medical service quality control and improvement center [Web page in Chinese]. Peking University Third Hospital. 2025. URL: https://www.puh3.net.cn/info/1971/21191.htm [accessed 2026-01-09]
33. Francis JJ, Johnston M, Robertson C, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health 2010 Dec;25(10):1229-1245. [doi: 10.1080/08870440903194015] [Medline: 20204937]
34. Talkhan H, Stewart D, McIntosh T, et al. Exploring determinants of antimicrobial prescribing behaviour using the Theoretical Domains Framework. Res Social Adm Pharm 2024 Apr;20(4):401-410. [doi: 10.1016/j.sapharm.2023.12.009] [Medline: 38320948]
35. Bele S, Cassidy C, Curran J, Johnson DW, Bailey JAM. Using the Theoretical Domains Framework to identify barriers and enablers to implementing a virtual tertiary-regional telemedicine rounding and consultation for kids (TRaC-K) model: qualitative study. J Med Internet Res 2021 Dec 22;23(12):e28610. [doi: 10.2196/28610] [Medline: 34941561]
36. 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):24047204. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
37. Hearn JH, Bryson K, Barsauskaite L, Bullo S. A COM-B and Theoretical Domains Framework mapping of the barriers and facilitators to effective communication and help-seeking among people with, or seeking a diagnosis of, endometriosis. J Health Commun 2024 Mar 3;29(3):174-186. [doi: 10.1080/10810730.2023.2299003] [Medline: 38262475]
38. Wang L, Wu T, Guo X, Zhang X, Li Y, Wang W. Exploring mHealth monitoring service acceptance from a service characteristics perspective. Electron Commer Res Appl 2018;30:159-168. [doi: 10.1016/j.elerap.2018.06.003]
39. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
40. Eyrich NW, Andino JJ, Fessell DP. Bridging the digital divide to avoid leaving the most vulnerable behind. JAMA Surg 2021 Aug 1;156(8):703-704. [doi: 10.1001/jamasurg.2021.1143] [Medline: 33769441]
41. 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. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
42. Yang Y, Zhu X, Song R, Zhang X, Guo F. Not just for the money? An examination of the motives behind physicians’ sharing of paid health information. J Inf Sci 2023 Feb;49(1):145-163. [doi: 10.1177/0165551521991029]
43. Haji FA, Rojas D, Childs R, de Ribaupierre S, Dubrowski A. Measuring cognitive load: performance, mental effort and simulation task complexity. Med Educ 2015 Aug;49(8):815-827. [doi: 10.1111/medu.12773] [Medline: 26152493]
44. Ansarian M, Baharlouei Z. Applications and challenges of telemedicine: privacy-preservation as a case study. Arch Iran Med 2023 Nov 1;26(11):654-661. [doi: 10.34172/aim.2023.96] [Medline: 38310426]
45. Churruca K, Foo D, Turner A, et al. Users’ perspectives of direct-to-consumer telemedicine services: survey study. JMIR Form Res 2025 Feb 3;9:e68619. [doi: 10.2196/68619] [Medline: 39899749]
46. Zhang X, Guo X, Guo F, Lai KH. Nonlinearities in personalization-privacy paradox in mHealth adoption: the mediating role of perceived usefulness and attitude. Technol Health Care 2014;22(4):515-529. [doi: 10.3233/THC-140811] [Medline: 24763205]
47. Choi N. Relationship between health service use and health information technology use among older adults: analysis of the US National Health Interview Survey. J Med Internet Res 2011 Apr 20;13(2):e33. [doi: 10.2196/jmir.1753] [Medline: 21752784]
48. Jing R, Lai X, Li L, Wang H. Internet use and healthcare utilization among older adults in China: a nationally representative cross-sectional study. BMC Geriatr 2025 Mar 5;25(1):150. [doi: 10.1186/s12877-025-05798-9] [Medline: 40045191]
49. Mishkin AD, Zabinski JS, Holt G, Appelbaum PS. Ensuring privacy in telemedicine: ethical and clinical challenges. J Telemed Telecare 2023 Apr;29(3):217-221. [doi: 10.1177/1357633X221134952] [Medline: 36349356]
50. Pool J, Akhlaghpour S, Fatehi F, Gray LC. Data privacy concerns and use of telehealth in the aged care context: an integrative review and research agenda. Int J Med Inform 2022 Apr;160:104707. [doi: 10.1016/j.ijmedinf.2022.104707] [Medline: 35131698]
51. Solimini R, Busardò FP, Gibelli F, Sirignano A, Ricci G. Ethical and legal challenges of telemedicine in the era of the COVID-19 pandemic. Medicina (Kaunas) 2021 Nov 30;57(12):1314. [doi: 10.3390/medicina57121314] [Medline: 34946259]
52. Sun L. Health data governance in China: emphasizing ‘sharing’ and ‘protection’ based on the right to health. Med Law Int 2023 Mar;23(1):26-43. [doi: 10.1177/09685332221140416]

53. Zhang H, Zhang H, Zhang Z, Wang Y. Patient privacy and autonomy: a comparative analysis of cases of ethical dilemmas in China and the United States. BMC Med Ethics 2021 Feb 2;22(1):8. [doi: 10.1186/s12910-021-00579-6] [Medline: 33531011]
54. Huang YN, Yin WQ, Liu Q, Wen QZ, Huang DM, Chen ZM, et al. Analysis on the relationship between residents' medical trust and social trust [in Chinese]. Chin Health Serv Manage 2018;35(10):777-779.
55. Mi J, Du X, Aili Y, Li M, Liu Y. Unveiling patient trust in physicians across China: insights from a nationwide cross-sectional study. Risk Manag Healthc Policy 2024;17:1039-1052. [doi: 10.2147/RMHP.S456585] [Medline: 38680478]

Abbreviations

mHealth: mobile health
TAM: technology acceptance model
TDF: Theoretical Domains Framework
UTAUT: unified theory of acceptance and use of technology

Edited by A Kushniruk; submitted 03.Jun.2025; peer-reviewed by J Claggett, X Zhang; revised version received 03.Dec.2025; accepted 28.Dec.2025; published 21.Jan.2026.

Please cite as:
Liu K, Yang Y, Song Z, Li H, Lyu Y, Zhang K, Wu X, Hou Z, Wang Y
Using the Theoretical Domains Framework to Identify Barriers to and Enablers of Patient Telemedicine Services Use in China: Qualitative Study
JMIR Hum Factors 2026;13:e78457
URL: https://humanfactors.jmir.org/2026/1/e78457
doi:10.2196/78457

Hospital-at-Home for South Asian Communities in British Columbia, Canada: Qualitative Interview Study

Emma Wong[1], MPH; Mahabhir Kandola[1], MSc; Kamal Arora[1], PhD; Harroop Sharda[2], PhD; Roman Deol[1], BSc, MBA; Mary Jung[3], PhD; Robert Paquin[4], BSN, MSc; Maria Montenegro[1], MSc; Megan MacPherson[1,5], PhD

[1]Department of Virtual Health, Fraser Health, 13450 102 Ave, Surrey, BC, Canada
[2]British Columbia Institute of Technology, Burnaby, BC, Canada
[3]School of Health and Exercise Sciences, University of British Columbia, Kelowna, BC, Canada
[4]King's College London, London, United Kingdom
[5]Department of Occupational Science and Occupational Therapy, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Megan MacPherson, PhD
Department of Virtual Health, Fraser Health, 13450 102 Ave, Surrey, BC, Canada

Abstract

Background: South Asian communities in Canada face significant disparities in access to health care and experience higher rates of chronic conditions such as cardiovascular disease, diabetes, and hypertension. Hospital-at-Home services have the potential to improve access and outcomes, yet little is known about how these services are perceived and experienced by South Asian patients and caregivers. Understanding both barriers and facilitators is critical for culturally responsive implementation.

Objective: This study aimed to explore the experiences of South Asian community members with in-person hospital care and their perceptions, attitudes, and expectations regarding virtual Hospital-at-Home services, with the goal of identifying culturally tailored strategies to improve access, quality, and satisfaction.

Methods: A qualitative study using semistructured interviews was conducted with 20 South Asian community members in the Fraser Health region in British Columbia, Canada. Interviews explored experiences with in-person hospital care, perceptions of a virtual hospital service (also known as Hospital-at-Home), and recommendations for enhancing awareness and accessibility. Interviews were audio-recorded, transcribed, and analyzed thematically to identify key patterns in perceptions, experiences, and needs.

Results: Participants described multiple systemic barriers to in-person hospital care, including long wait times, overcrowding, transportation challenges, and difficulty navigating the health system. Cultural and religious needs, such as gender-concordant care and culturally appropriate food, were frequently unmet, while language-concordant care and family involvement were critical to positive experiences. Discrimination and assumptions based on ethnicity or age further shaped perceptions of care. Virtual hospital services were valued for convenience, comfort, reduced exposure to hospital-acquired infections, and support for family involvement. However, participants raised concerns about clinical quality, the absence of physical examinations, digital literacy, privacy, and home-based responsibilities. Acceptance varied by age, immigration status, and familiarity with technology. Participants emphasized the importance of culturally tailored outreach, leveraging community leaders, ethnic media, and peer testimonials to increase awareness and trust.

Conclusions: South Asian patients and caregivers recognize both challenges in traditional hospital care and potential benefits of Hospital-at-Home services. Implementation strategies that address systemic barriers, integrate cultural and linguistic considerations, and engage trusted community networks are essential to improving equity, access, and satisfaction. Findings highlight the need for culturally responsive, patient-centered approaches in the design and delivery of virtual health services for racialized populations.

(JMIR Hum Factors 2026;13:e79675) doi:10.2196/79675

KEYWORDS

telemedicine; South Asian people; qualitative research; health equity; Hospital-at-Home

Introduction

Background

The South Asian population, which includes individuals with ethnic roots originating from Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka, constitutes the largest visible minority or racialized group in Canada [1]. South Asians in Canada face a higher prevalence of cardiovascular disease, diabetes, and hypertension compared to other racialized groups [2,3]. While research suggests that genetic predisposition is a partial contributor [4], the observed health disparities likely also result from social and environmental determinants of health, such as access to higher education, employment opportunities, food security, experiences of discrimination, and access to culturally appropriate health care services [5-7]. It is important to recognize that South Asian populations in Canada are highly diverse, encompassing a range of languages, religions, socioeconomic backgrounds, and migration histories, all of which influence health experiences and outcomes. Previous research has shown that South Asians encounter significant barriers in accessing health care resources for chronic disease prevention and management [8]. Given the inequities and disparities faced by South Asian community members within Canada, innovative care strategies are needed to enhance access to culturally relevant care and improve health outcomes.

Virtual Health Care

Virtual health care, which uses remote technologies to deliver health services, has been identified as a key mechanism for improving access to care [9]. Specifically, virtual hospitals-at-home (commonly referred to as simply Hospital-at-Home) provide an alternative to traditional “brick and mortar” hospital admissions, allowing patients to receive the care they need from the comfort of their own home [10]. This care model has been increasingly used worldwide, accelerated by hospital capacity strains during the COVID-19 pandemic [11,12]. While admitted to a Hospital-at-Home, patients are monitored remotely using digital devices that transmit data to their multidisciplinary care team [10] and have daily virtual consultations via telephone or video, supplemented with home visits when necessary [10].

Research has shown that Hospital-at-Home models are both feasible and effective across various conditions [13]. Studies report that patients opt into virtual hospitals at high rates, and that this approach delivers high-quality care with fewer complications (eg, shorter hospital stays, lower incidence of delirium, and reduced rate of falls [14]). Patients and caregivers frequently express higher satisfaction with home-based care, citing increased comfort, privacy, and family engagement as key benefits [13].

Beyond individual benefits, Hospital-at-Home can alleviate system-wide pressures by reducing wait times, freeing up hospital bed capacity, and expanding access to care [15]. Despite these advantages, challenges remain. Some studies have noted higher readmission risk for patients with complex medical conditions, as well as increased caregiver burden in home settings [13]. Additionally, concerns about communication barriers and the need for structured care coordination highlight areas where further improvements are needed to ensure equitable and effective implementation of Hospital-at-Home [13].

Given the potential of virtual hospitals to enhance access to care, particularly for underserved populations, it is critical to understand and address the barriers to their adoption and use. This is especially important for racialized communities, such as South Asians in Canada, who already experience health disparities and face unique challenges in accessing virtual health services [16].

Virtual Health Use in South Asian Communities

While virtual care has the potential to alleviate barriers experienced by South Asian communities [17] in Canada, and research suggests positive attitudes toward integrating virtual services [16], significant barriers still exist for accessing these services. Studies have shown that age, gender, income, education, and language are associated with the ability to use health-related technologies among Panjabi-speaking (also known as “Punjabi”) South Asian individuals in Canada [18] and that language preferences, education, age, and sex predict the use of virtual care among South Asians in Alberta [19]. In both studies, older adults, women, those with a preference for non-English communication, and those with lower educational levels were less likely to engage with virtual care [18,19]. A recent community-based participatory action research project in Surrey, British Columbia, Canada further supported these findings [16]. Through focus groups and photo voice activities, participants identified concerns about their ability to navigate new technologies as they age, gendered responsibilities limiting women’s access to technology, and communication difficulties due to language barriers, literacy levels, and digital literacy [16].

Although virtual care is designed to support disease prevention and self-management, these services are often developed without equity in mind, are not culturally responsive, and are generally underused by South Asian populations in Canada [16]. The divide in access and engagement with virtual care services could further exacerbate health disparities faced by South Asian communities [8]. Given the relative novelty of virtual hospitals, little is known about their acceptance among South Asian communities. Future virtual care models must address known barriers to ensure equitable access.

Setting

Health care in British Columbia, Canada, is provided by a provincial health authority, 5 regional health authorities, and a First Nations health authority [20]. The Fraser Health Authority is the largest regional health authority, serving over 2 million people across 20 diverse and rapidly growing communities [21]. According to the 2021 Census, nearly 1 in 5 Fraser Health residents identifies as South Asian [22]. Fraser Health emphasizes understanding how religion, language, and cultural practices influence health to develop customized programs and services.

To better support South Asian communities, Fraser Health established the South Asian Health Institute (SAHI) in 2013, focusing on research, innovation, and evidence-based care [5]. Fraser Health’s Virtual Health team supports the development and implementation of new virtual care pathways across the region. One such pathway is the virtual hospital service. As Fraser Health serves 75% of South Asian residents in the province, it is essential to conduct patient-oriented research, or research done in direct partnership with patients, to understand what matters to South Asian communities when developing the virtual hospital [5].

Research Objectives

This patient-oriented research study aims to engage South Asian community members through 3 primary objectives:

1. Capture and understand the emotional experiences, as well as the perceived barriers and facilitators, of members within South Asian communities in the context of both in-person and Hospital-at-Home systems.
2. Understand the lived experiences of South Asian community members in navigating the health care system and identify gaps in accessing care.
3. Identify opportunities for how Hospital-at-Home services can be implemented to improve the access and flow of care among South Asian communities.

Methods

Overview

This qualitative study used semistructured interviews to explore the experiences, perceptions, and attitudes of South Asian community members in British Columbia, Canada, toward in-person hospital care and the emerging Virtual Hospital-at-Home model. The study was guided by a patient-oriented research approach, prioritizing the perspectives of patients and caregivers to inform culturally relevant virtual care strategies.

This study was guided by a steering group including members from Progressive Intercultural Community Services (PICS) Society [23], Fraser Health's Virtual Health department [24], Fraser Health's SAHI [25], and Fraser Health's Equity, Diversity, and Inclusion team [26], and academic partners from the University of British Columbia and King's College London.

Interviews were guided by a pragmatic epistemology, which seeks to generate practical findings that can be readily applied in real-world settings while also tolerating multiple truths [27]. Although all South Asian community members may be exposed to the same care environments, everyone has unique experiences.

Recruitment

Participants were recruited through convenience sampling using outreach by the PICS Society and the SAHI, and supplemented with community outreach efforts including social media, places of worship, and public libraries. The inclusion criteria were that participants must (1) identify as South Asian, (2) reside within the Fraser Health region, and (3) have been a patient in the hospital in the past year or have been a caregiver to someone in the hospital in the past year. Recruitment continued until thematic saturation was reached, when no new concepts emerged from subsequent interviews.

Data Collection

South Asian community members were interviewed to understand their lived experiences navigating the virtual health care system. A topic guide, adapted from previous Virtual Health patient partner workshops, was revised with feedback from community members and edited by Fraser Health's plain language team for clarity. The guide covered 2 primary domains: (1) experiences with in-person hospital care, including barriers, facilitators, and family involvement; and (2) perceptions of, and attitudes toward, the Virtual Hospital-at-Home model, including benefits, concerns, and strategies for community engagement. Interviews began with a screening question to classify participants as patients, caregivers, or both. A full interview topic guide is provided in Multimedia Appendix 1.

Thirty- to 60-minute, one-on-one interviews were conducted via Microsoft Teams, with the PICS Society providing technical support when needed to ensure participant access to the platform. To ensure accessibility, interpretation services were available in Panjabi, Hindi, Urdu, and Pashto. All interviews were audio-recorded with participant consent, transcribed verbatim, and translated where required by members of the research team. Interviews were conducted by EW and MK, who both had master's degrees and worked as a "Research Assistant" and the "Research and Knowledge Translation Lead" for the Virtual Health department at the time of interviews. All interviewees had no prior relationships with those conducting the interviews.

Data Analysis

Thematic analysis guided by the approach of Braun and Clarke [28] was conducted using NVivo (QSR International). First, transcripts were read repeatedly for familiarization. An initial set of open codes was inductively generated to capture salient features of the data. Codes were iteratively refined and grouped into higher-order categories. Themes were then developed through constant comparison across transcripts, paying attention to both convergent and divergent perspectives. Theme prevalence was assessed approximately by noting how many participants raised each theme, though exact tallies were not feasible given the conversational variability of interviews. Findings were also interpreted with attention to cultural safety and equity considerations relevant to South Asian communities in the Fraser Health region.

Coding was conducted by MK and cross-checked by MM; discrepancies were resolved by consensus. For multilingual interviews, translations were reviewed by bilingual team members to ensure conceptual equivalence. Final themes were reviewed and refined in consultation with the steering committee to enhance cultural validity. The steering committee did not influence the coding of raw data. Thematic saturation was achieved by the 17th interview; 3 additional interviews confirmed that no new themes emerged.

Ethical Considerations

This study received approval from the Fraser Health Research Ethics Board (H23-03144). Interested participants were sent a consent form, the topic guide, and a video explaining the Hospital-at-Home to review beforehand. Informed consent for participation and interview recording was obtained verbally prior to the interview. To protect privacy and confidentiality, all transcripts were anonymized, and identifiers were removed prior to analysis. Data were stored securely on password-protected institutional servers. Participants were compensated with a US $14.86 gift card at the conclusion of their interview.

Results

Participant Characteristics

Twenty participants were interviewed (n=9 male, n=11 female). Participants were aged 19 - 79 years, most commonly between 60 - 69 years (6/20, 30%). Younger adults aged 19 - 29 (n=5), 30 - 39 (n=4), and 40 - 49 (n=4) years were also represented, along with 1 participant in their 70s. Interviews were conducted in English (n=15), Panjabi (n=3), Pashto (n=1), and Hindi (n=1).

Most participants were originally from India (14/20, 70%), with others from Afghanistan (3/20, 15%), and smaller representations from Iraq (n=1), Pakistan (n=1), and India and Fiji (n=1). Half had lived in the community for more than 10 years (n=10); the remainder were more recent arrivals (n=4 lived in Canada <1 year, n=5 lived here 1 - 3 years). Regarding immigration status, 9 were Canadian citizens and 11 were permanent residents.

Self-rated health was generally positive—6 participants (30%) rated their health as “very good,” 6 (30%) as “good,” and 8 (40%) as “average.” Six participants (30%) reported having a chronic condition. Participants included caregivers only (7/20, 35%), patients only (6/20, 30%), and both caregivers and patients (5/20, 25%). Two participants also contributed professional perspectives, working in newcomer services and health care.

All participants had access to a smartphone (20/20, 100%), while 11 (55%) had access to a tablet and 11 (55%) had access to a laptop. Ten participants (50%) had all 3 devices, while 8 (40%) had only a smartphone. Device use was frequent—14 (70%) used devices daily or multiple times per day, 4 (20%) used devices regularly but not daily, and 2 (10%) did not regularly use devices.

Experiences of In-Person Hospital Care

Overview

Participants described a mix of challenges and positive experiences with in-person hospital care. Six overarching themes emerged: systemic barriers to accessing care, communication and language barriers, cultural and religious needs not met, family and caregiver roles, discrimination and bias, and positive experiences and gratitude. Experiences varied by age, gender, immigration status, and length of time in Canada. Older and more recently arrived participants reported more barriers than younger or Canadian-born participants.

Systemic Barriers to Accessing Care

Participants frequently reported challenges accessing timely and appropriate hospital services, including long wait times, overcrowding, transportation difficulties, and navigation challenges. These barriers contributed to frustration, fatigue, and inequitable access.

Long Wait Times

Wait times were among the most common challenges mentioned. Many participants described hours-long waits in emergency departments, sometimes extending overnight or across multiple days. As 1 participant shared, “we have waited with a broken wrist 28 hours” (P10, female, aged 60 - 69 years, >10 years in community). Another recalled, “you have to wait maybe sometimes 10 hours or more than that and you have a critical illness” (P20, female, aged 60 - 69 years, >10 years in community). Such prolonged waits were perceived as not only inconvenient but also unsafe when health concerns were urgent.

Overcrowding and Understaffing

Participants described hospitals as overcrowded and short-staffed, with patients left in hallways and nurses visibly stretched thin. One participant reported, “patients are all outside in the hallways. They don’t even have proper rooms” (P13, female, aged 60 - 69 years, <1 year in community). Another observed, “sometimes you can see like the nurses are really tired, and then you see them overwhelmed with the amount of work” (P11, female, aged 19 - 29 years, >10 years in community). These accounts reflect how structural resource constraints were felt by both patients and providers.

Transportation Difficulties

Transportation presented another layer of inequity, particularly for patients without private vehicles. Some relied on long bus rides with multiple transfers, which became especially challenging for older adults and those with mobility issues. One participant explained,

> *The nearest hospital is like a 45 minute transit ride, but the bus doesn’t go directly there. So I have to walk. For me as an able-bodied male, mid-20s, it’s fine. But...for my grandma who doesn’t drive and can’t speak much English, it’s very hard.* [P1, male, aged 20 - 29 years, >10 years in community]

Another recalled, “I rely on HandyDART, and if that’s not available then my son would have to take time off work to drive me” (P17, female, aged 30 - 39 years, >10 years in community). These logistical challenges shaped how and when participants sought hospital care.

Navigation Challenges for Newcomers

For newcomers to Canada, navigating the health system added further barriers. Several participants spoke about confusion around referrals, insurance, and entitlements. One reflected,

> *I took around a year’s time to understand what was MSP and how to access it. What is covered under it and what do I need other insurance for?* [P8, female, aged 30 - 39 years, 1 - 3 years in community]

Another highlighted misunderstanding about referral pathways:

> *You have to first go to a family doctor, and then if the problem is serious then you go to hospital. Many of the people don’t have that understanding.* [P6, male, aged 40 - 49 years, >10 years in community]

Such experiences suggest that lack of system literacy can delay or complicate access to hospital care.

Language and Communication

Communication challenges were central to participants' accounts, particularly for older adults and newcomers with limited English proficiency, with many describing difficulties conveying symptoms, understanding instructions, or navigating interactions without the help of family. While some praised interpreters or language-concordant providers, gaps in communication were widely felt and often shaped the overall hospital experience.

Limited English Proficiency

Participants often highlighted the difficulty of explaining their health concerns in English. For example, one explained,

> *If I talk about my grandfather, they don't know how to speak in English...if they are taking them alone into the rooms, they can't express how they're feeling or they can't tell about their problems.* [P5, male, aged 19 - 29 years, 1 - 3 years in community]

Another noted, "when the other person is not from my community, the only hurdle I feel is the language" (P9, female, aged 30 - 39 years, 4 - 6 years in community). These challenges not only caused stress but also risked miscommunication in clinical encounters and increased reliance on family members acting as interpreters. One participant described, "she had a family member of hers that spoke on her behalf and communicated with the doctors" (P4, female, aged 60 - 69 years, >10 years in community). Similarly, another shared, "mom didn't speak English, so until I got there and started helping her she didn't have any good idea as to what was happening" (P14, male, aged 60 - 69 years, >10 years in community). While this reliance on family interpreters ensured patients could understand their care, it also created dependence on family members, sometimes leading to delays or added burdens. Participants also noticed changes over time, including greater availability of translators and culturally diverse staff. As one shared, "now they have translators everywhere, even in the parking. Greeters in the hospital who speak their language" (P17, female, aged 30 - 39 years, >10 years in community). Another noted, "everywhere there are translators in hospitals" (P18, male, aged 70 - 79 years, >10 years in community).

Language-Concordant Care

Participants expressed strong appreciation for health care staff who spoke their language, emphasizing that this reduced stress and improved trust. As one noted, "we prefer to go to [specific hospital] just because the nurses and all the other staff member do speak Panjabi. So it's easier for my grandparents" (P11, female, aged 19 - 29 years, >10 years in community). Another explained, "if the person is from South Asian communities and they are feeling shy or hesitating to say something in front of a White person, a South Asian doctor or nurse can help" (P5, male, aged 19 - 29 years, 1 - 3 years in community). Such accounts highlight how language-concordant care fosters both clarity and cultural comfort.

Cultural and Religious Needs Not Met

Participants frequently described hospital environments that did not align with their cultural or religious values. These included concerns about gender concordance with providers, lack of culturally appropriate food, and perceived insensitivity to religious practices.

Gender-Concordant Care

Gender of providers was an important factor for some participants, particularly for women from Muslim and Sikh backgrounds. One explained, "for me as a Muslim lady, I don't want anybody to touch me if he's a man and get naked in front of him while there is another lady available" (P20, female, aged 60 - 69 years, >10 years in community). Another recalled, "she would have preferred a female nurse, but they had a male" (P1, male, aged 19 - 29 years, >10 years in community). These accounts underscore how gender concordance shaped comfort and dignity in care.

Dietary and Religious Respect

Food was another area where participants felt their needs were not adequately met. Some reported being served meals that conflicted with their dietary or religious practices. For instance, 1 participant explained, "she was a very strict vegetarian, but they fed her beef" (P1, male, aged 19 - 29 years, >10 years in community). Others described hospital food as unappealing or culturally inappropriate:

> *It's a challenge because the hospital food is a challenge for them. They don't like it, obviously, because it's not their cultural food.* [P17, female, aged 30 - 39 years, >10 years in community]

Cultural Insensitivity

In some cases, participants felt their cultural or religious practices were disrespected. One recalled, "she had her religious books and statues by her bed and they moved to a corner by garbage, which is offensive to our religious practices" (P1, male, aged 19 - 29 years, >10 years in community). These incidents, though not universal, left lasting negative impressions and diminished participants' sense of respect for care. Participants recounted instances where staff failed to accommodate dietary and religious needs, sometimes in ways that were profoundly disrespectful.

Family and Caregiver Roles

Family members were described as essential to the hospital experience, serving as translators, advocates, and sources of emotional support. Yet participants also reported that hospital policies and circumstances sometimes limited family involvement, creating both barriers and additional burdens.

Restrictions on Family Presence

Some participants described frustration with restrictions on family members accompanying patients, even when needed for support or translation. One participant recalled, "they were not allowing my family members to come and sit with me" (P2, male, aged 19 - 29 years, <1 year in community). Another shared, "sometimes the person needs emotional support or they need someone from the family...but I can't go with her because they don't have enough space" (P5, male, aged 19 - 29 years, 1 - 3 years in community). Such restrictions were perceived as leaving patients vulnerable and isolated.

Emotional and Practical Burden on Families

While family involvement was crucial, it often created strain. One participant described, "her son would take her, but he would have to take time off work in order to take her" (P4, female, aged 60 - 69 years, >10 years in community). Another reflected on balancing childcare with caregiving: "sometimes I have to be without food for long times...I cannot also go to the cafeteria with kids" (P9, female, aged 30 - 39 years, 4 - 6 years in community). These stories illustrate how the health system's reliance on families carries significant costs.

Preference for Family Caregiving

While family involvement can be burdensome, many participants preferred it over professional care for trust and cultural reasons. As 1 participant explained, "South Asian people...are more comfortable with their family member, rather than a nurse, doing certain tasks" (P17, female, aged 30 - 39 years, >10 years in community). This highlights how family involvement was not only necessary but often desired.

Discrimination and Bias

Although many participants spoke positively of their care, some described experiences of stereotyping, dismissal, or unequal treatment.

Stereotyping and Assumptions

A few participants reported feeling judged on the basis of cultural stereotypes. One explained, "because they're Indian, they obviously eat unhealthy and because of that they have high blood pressure and heart issues...without doing further research" (P1, male, aged 19 - 29 years, >10 years in community). Such assumptions left participants feeling unseen and unfairly blamed.

Unequal Quality of Care

Others felt that seniors and immigrants received less attentive or less culturally appropriate care. As one participant described, "I've seen that personally with my grandma, that they don't take seniors that seriously" (P11, female, aged 19 - 29 years, >10 years in community). Another reflected, "it's not very culturally appropriate...my grandparents' experience is not the same" (P1, male, aged 19 - 29 years, >10 years in community). These perceptions shaped trust in the system and influenced decisions about where and when to seek care.

Positive Experiences and Gratitude

Despite systemic and cultural challenges, participants consistently recognized skilled, compassionate care. One said,

> *The staff was very good. They were peacefully talking. The doctor was very nice. It was very good.* [P18, male, aged 70 - 79 years,>10 years in community]

Another reflected, "I have the best experience and I was very happy" (P3, male, aged 40 - 49 years, 1 - 3 years in community). Additionally, trust in providers' skill was often expressed as gratitude, even when other barriers were present. For instance, 1 parent stated, "I'm very grateful to all the doctors and nurses who provided care to my child" (P12, female, aged 40 - 49 years, <1 year in community). Such reflections reveal the duality of participants' experiences: systemic and cultural barriers coexisted with respect for the expertise and dedication of individual providers.

Perceptions and Expectations of Hospital-at-Home

Overview

Participants shared diverse perspectives on virtual hospital services, highlighting benefits, limitations, and conditions under which such care is acceptable. Five major themes emerged—convenience, comfort, and family support; cultural and generational differences; quality of care and clinical limitations; emotional safety, privacy, and trust; and barriers to access and technology. Participants' experiences and expectations reflected both practical considerations, such as travel and technology, and cultural and interpersonal factors influencing acceptance. Views often diverged across age, gender, and immigration status, with younger or Canadian-born participants more enthusiastic about virtual care and elders or newcomers expressing hesitation.

Convenience, Comfort, and Family Support

Overview

Participants consistently highlighted the practical and emotional advantages of receiving care at home. Virtual hospital services were valued for reducing stress, increasing comfort, and enabling family involvement, which was seen as supportive both physically and emotionally. Many participants described how being in a familiar environment could enhance recovery, while also reducing exposure to illness compared with hospital settings.

Avoiding Travel and Wait Times

Many participants appreciated the time-saving and stress-reducing aspects of virtual care. One noted,

> *I can be at home and go about my daily things and just wait for a phone call. I don't have to actually sit at all at an actual office anyway.* [P6, male, aged 40 - 49 years, >10 years in community]

Others highlighted avoidance of transportation challenges and hospital stress, stating it would allow them "to get help like at home...I don't have to stuck in the traffic and get stressed out" (P10, female, aged 60 - 69 years, >10 years in community) and "virtual hospitals would be more convenient. He won't have to worry about transportation waiting. Time to come and go" (P7, male, aged 60 - 69 years, 1 - 3 years in community).

Healing at Home

Being at home with family and in a familiar environment was viewed as beneficial for recovery. Participants described emotional and physical advantages: "you are around your family, that motivates you to be more good. I think that's the biggest thing that virtual can provide" (P2, male, aged 19 - 29 years, <1 year in community), and "if this virtual hospital help was available, then definitely he would have felt more better at home rather than at the hospital...we can read books for him, we can sit beside him" (P17, female, aged 30 - 39 years, >10 years in community). Others emphasized comfort and food preferences, noting,

> *I feel more comfortable in my space...Benefits I can be with my family. That’s the best benefit I can eat the food that I want.* [P16, female, aged 30 - 39 years, <1 year in community]

Reduced Exposure to Illness

Participants recognized the safety advantage of avoiding hospital-acquired infections: “It can help many people...rushing to the hospitals because of. danger of getting communicable diseases” (P12, female, aged 40 - 49 years, <1 year in community), and “with the virtual, then there wouldn’t be so many people waiting for beds. hospitals have a lot of diseases going around” (P15, female, aged 19 - 29 years, >10 years in community).

Smooth Transitions from Hospital to Home

Several participants highlighted the importance of smooth transitions from hospital to home, emphasizing clear communication, rapid response in emergencies, and reliable monitoring. The availability of equipment and the teach-back style training were viewed as critical.

> *...how the South Asian will be educated...equipment will be there, they’ll teach you how to use it. It’s important that you use it properly.* [P4, female, aged 60 - 69 years, >10 years in community]
>
> *If something gets worse, who do I call first? That should be very clear.* [P15, female, aged 19 - 29 years, >10 years in community]

Home-Based Challenges

Participants noted that home environments can pose challenges during recovery, such as stairs, inaccessible bathrooms, or absence of hospital equipment:

> *My bedroom is upstairs and the bathroom is downstairs. After surgery, how can I manage the stairs?* [P7, male, aged 60 - 69 years, 1 - 3 years in community]
>
> *We don’t have special hospital beds at home. Sometimes you need that equipment to recover properly.* [P12, female, aged 40 - 49 years, <1 year in community]

Routine household responsibilities, such as childcare, also limited the ability to rest: “I know if I go home, I have to cook for the other two children and I have to take care of them” (P20, female, aged 60 - 69 years, >10 years in community). Participants worried that home-based care could shift responsibilities onto family members, especially women and elders:

> *Yes, absolutely, definitely. When I am at the hospital, I’m just quite confident that I’ve been taken care of 100%. But if I am at home, it will put more pressure on me* [P10, female, aged 60 - 69 years, >10 years in community]

Cultural Norms and Generational Differences

Overview

Cultural norms, family expectations, and generational comfort with technology strongly influenced participants’ perceptions of virtual hospital care. While younger participants reported ease with online platforms, older adults often struggled with technology and valued in-person monitoring. Cultural beliefs regarding family roles and appropriate care also shaped attitudes toward remote health care.

Generational Divide in Comfort with Technology

Participants highlighted that younger generations were more comfortable with technology, while older adults often struggled. One reflected,

> *For me, growing up in this generation, I’m not shy for technology, but for example, my grandparents, it’s very a new world to them. So for them there is more of that disconnect.* [P1, male, aged 19 - 29 years, >10 years in community]

Similarly, a participant noted, “if I imagine my grandmother using this service, she might be overwhelmed with the tablet thing...she would just be like, I just wanna meet people in real” (P16, female, aged 30 - 39 years, <1 year in community).

Cultural Norms and Expectations of Care

Some participants described South Asian cultural preferences for in-hospital monitoring, emphasizing family support and safety:

> *I believe for the South Asian people this will not work very well...I would say at the hospital, I feel the care is around me 24 hours...So I think at the hospital we can get the maximum care.* [P20, female, aged 60 - 69 years, >10 years in community]

Others stressed the importance of culturally sensitive approaches, including language support: “I think having that community-centered culturally sensitive approach to healthcare is really important because this generic advice doesn’t help the community*”* (P1, male, aged 19 - 29 years, >10 years in community).

Absence of Support Persons at Home and Immigrant Social Networks

Some participants identified the absence of support persons at home as a barrier to virtual care. As immigrants, many noted that they no longer have the robust support systems they relied on in their native countries. While they can manage their daily lives, participants expressed doubts about their ability to maintain the same level of self-care when unwell, particularly those living alone or with family members who work full time.

> *It’s very hard because you don’t have much people around you...You don’t have a community...You don’t have neighbors as we use to have uh back in in Asia or South East Asia, there’s a lot of people around you...But here, because we are immigrant, it’s very limited people that we know and it’s not the bond even it’s not that because of the type of living here as*

> *we all work or just us to take care of the family...So our relationship is very limited.* [P20, female, aged 60 - 69 years, >10 years in community]

Quality of Care and Clinical Limitations

Overview

While participants acknowledged the convenience of virtual care, many expressed concerns about clinical quality. The inability to conduct physical examinations and the perception that remote care could be “lesser” than in-person attention were recurring issues. Participants suggested virtual care might be suitable for routine or minor issues, but not for serious medical concerns.

Value of Physical Examination

Many participants expressed concern that virtual care could not substitute for hands-on assessment:

> *In my view...the Virtual Hospital will not be very effective if the patient is not present in front of the doctor...virtual care will be only communication.* [P3, male, aged 40 - 49 years, 1 - 3 years in community]
>
> *My knee problem...I wasn’t sure how I could show my knees over the phone?* [P13, female, aged 60 - 69 years, <1 year in community]

Perceptions of “Lesser Care”

Some participants worried that virtual hospital services could be perceived as inferior, noting, “I just feel like if it’s online like people would think like, oh, you’re not taking me seriously” (P11, female, aged 19 - 29 years, >10 years in community) and “they wouldn’t fully be supportive...they would think ohh I would receive better care at a hospital rather than me taking my medications on my own” (P15, female, aged 19 - 29 years, >10 years in community). Participants suggested virtual care is appropriate for routine or minor concerns, while serious conditions require in-person attention:

> *If it’s a small thing, then Virtual Hospital is better. But if it’s like a big issue, then she would prefer in-person care.* [P4, female, aged 60 - 69 years, >10 years in community]

Emotional Safety, Privacy, and Trust

Overview

Emotional well-being, privacy, and trust emerged as key considerations. Participants emphasized the importance of empathy and interpersonal connection, while also noting concerns about privacy at home and the potential for fraud or security issues.

Need for Interpersonal Connection

Participants emphasized empathy and human touch, even via remote communication:

> *Having that interpersonal touch as much as you can...sometimes having that personal touch makes them feel better. But through a phone, you don’t really get that personal touch, so still making sure whoever is on the line is caring and kind.* [P1, male, aged 19 - 29 years, >10 years in community]

Concerns About Privacy at Home

Some expressed difficulties discussing sensitive issues at home:

> *If I want to tell you something, but I don’t want to tell other people at home it could be a concern. So it’s better to go to Doctor [in person].* [P10, female, aged 60 - 69 years, >10 years in community]

Trust in Fraser Health’s Professionalism

Participants generally assumed privacy and security were managed by a reputable provider: “I would also assume because Fraser Health is a reputed professional organization, they would have already taken all of these requisite steps to make sure their privacy is covered” (P6, male, aged 40 - 49 years, >10 years in community), and “I feel like Fraser Health has really good system for privacy” (P11, female, aged 19 - 29 years, >10 years in community).

Apprehension About Fraud or Scams

Some participants noted the risk of unknown numbers or digital communication: “Only thing is sometimes when you get a call from an unknown number, it seems like it might be a fraud” (P7, male, aged 60 - 69 years, 1 - 3 years in community).

Barriers to Access and Technology

Despite interest in virtual care, participants noted practical barriers related to technology and digital literacy. These challenges were particularly salient for older adults and recent immigrants.

Digital Literacy Challenges

Older adults may lack experience with technology: “Maybe having like info sessions at temples. helping teach the elders about how it works” (P1, male, aged 19 - 29 years, >10 years in community), and “if I imagine my grandmother using this service, she might be overwhelmed with the tablet thing” (P16, female, aged 30 - 39 years, <1 year in community).

Access to Equipment

Not all households have necessary devices: “Not just like assuming like they have a laptop on them...my mom’s mom has never touched a computer in her life” (P1, male, aged 19 - 29 years, >10 years in community).

Suggestions for Building Awareness of Virtual Hospitals at Home

Participants consistently highlighted the importance of proactive outreach and awareness-building, noting that many newcomers and elders were unaware of existing programs. Outreach strategies were rooted in community-based trust networks, emphasizing cultural familiarity, word of mouth, and visible engagement in South Asian gathering spaces.

Leveraging Trusted Spaces and Leaders

Participants recommended outreach through temples, gurdwaras, and mosques, where health authority representatives could host training or information sessions. They stressed that engagement would be more effective if facilitated by respected community leaders.

> *...meeting with the community leaders and they can like bring it to the community. I know at one of my temples...they have community members teaching them...maybe having someone from Fraser Health going in and doing a training session.* [P6, male, aged 40 - 49 years, >10 years in community]

Testimonials and Peer Influence

Most participants expressed hesitation about being the first to try the new program, preferring to hear positive testimonials from others before committing. They also mentioned that even a few negative reviews could deter their interest in the program. This reflects the strong sense of mutual trust and reliance on community members, which participants valued just as much as official information provided by the health authority. Hearing positive testimonials from peers was seen as essential to building trust and encouraging adoption.

> *...having like some kind of testimonial or proven proof that it is already worked before...that would just remove the stress or the doubt. Once it becomes like accepted as a normal thing, that would make it easy.* [P7, male, aged 60 - 69 years, 1 - 3 years in community]

Community Events and Visible Outreach

To engage South Asian communities, participants recommended venues such as libraries, community centers, places of worship, and networking events where community members regularly gather. Additionally, cultural and religious events, such as the annual Vaisakhi Day Parade held in Surrey, were identified as valuable outreach opportunities. Participants suggested setting up information booths equipped with printed materials, visual media, and representatives from the health authority to disseminate information about virtual care in the languages spoken by the community. They noted that organizations providing employment or settlement support (such as PICS) already use these channels effectively, demonstrating their potential for outreach. Furthermore, participants recommended collaborating with local South Asian–based nonprofit organizations.

> *...we have like festivals of the year, you know Vaisakhi, Culture Fest...maybe having like a Fraser Health pop up so community members can engage with it and educate it that way. Trying to get in the community where you can would be the best.* [P6, male, aged 40 - 49 years, >10 years in community]

Ethnic Media and Social Platforms

To reach a broader demographic, participants suggested using local South Asian newspapers, radio, or TV networks for individuals to share their lived experiences. This would particularly target the working population who often listen to the radio while driving. Additionally, some participants recommended using social media, a popular platform for news consumption, to further raise awareness.

> *I think doing more awareness about this virtual...talking with the people, providing more awareness to the people that they should know that such kind of program is existed.* [P7, male, aged 60 - 69 years, 1 - 3 years in community]

Discussion

Principal Findings

This qualitative study explored the experiences of South Asian community members in the Fraser Health region with in-person hospital services, and their perceptions and expectations of Hospital-at-Home models. Participants described significant systemic and cultural barriers in traditional hospital settings, alongside gratitude for providers’ skill and compassion. When discussing Hospital-at-Home, participants articulated both enthusiasm for its convenience and comfort and concerns about quality of care, cultural alignment, and feasibility in their home environments. Taken together, these findings provide important insights for health system planners seeking to advance culturally safe and equitable virtual hospital services.

Summary of Key Findings

Participants’ hospital experiences reflected 6 key themes: (1) systemic barriers, including long wait times, overcrowding, transportation, and navigation challenges; (2) communication and language barriers; (3) unmet cultural and religious needs; (4) essential but burdensome family and caregiver roles; (5) experiences of discrimination and bias; and (6) positive encounters characterized by gratitude and respect. These issues varied across age, gender, and immigration status, with older and more recently arrived participants reporting more barriers.

Expectations of Hospital-at-Home revealed 5 major themes: (1) convenience, comfort, and family support; (2) cultural and generational differences; (3) quality of care and clinical limitations; (4) emotional safety, privacy, and trust; and (5) barriers to access and technology. Across both sets of findings, participants consistently highlighted the central role of family, the importance of cultural sensitivity, and the need for accessible information delivered through trusted community channels.

Participants’ mixed experiences with in-person hospitals provide essential context for understanding perceptions of Hospital-at-Home. Structural challenges, such as prolonged emergency department waits and overcrowded wards, reduced trust in the timeliness and equity of hospital care. For newcomers, health system literacy gaps compounded these difficulties, as participants struggled to understand coverage, referral pathways, and navigation. These frustrations may increase openness to Hospital-at-Home as an alternative, particularly if it can reduce wait times and enhance comfort. However, systemic inequities that disadvantage newcomers and elders risk being reproduced in virtual formats unless addressed proactively.

Comparison to Existing Literature

Previous research has extensively documented barriers to virtual health care access for South Asians. Studies by Zibrik et al [18] and Makowsky et al [19] highlight disparities in virtual care access, while Dahal et al [29] and Hyman et al [16] emphasize the impact of socioeconomic factors on virtual health engagement. Notably, few studies have examined Hospital-at-Home programs within other racial or ethnic minority communities, limiting opportunities for direct comparison. This scarcity of literature underscores the novelty of exploring Hospital-at-Home acceptability specifically among South Asian populations in Canada and highlights the unique structural, cultural, and familial considerations identified in this study.

Findings from this study align with existing literature in confirming that language barriers, trust in health care providers, and digital literacy shape perceptions toward adopting virtual care. Moreover, consistent with Bhalla et al [30], community-driven strategies, such as the role of family members in health decision-making and the influence of social networks such as WhatsApp, play a crucial role in health care engagement for the South Asian community. These findings were echoed by participants of this study who noted their attitudes toward Hospital-at-Home care could be highly influenced by peer testimonials.

A key distinction of this study is its focus on Hospital-at-Home models rather than primary care or outpatient virtual services, which have been the emphasis of much prior research. While previous research has explored South Asians' engagement with digital tools for chronic disease management and primary care, this study surfaces concerns that are specific to the delivery of hospital-level care in the home. In particular, it contributes new insights on the role of the home environment, including the physical suitability of the space, the availability and involvement of family caregivers, the complexities of intergenerational living arrangements, and the burden of household responsibilities during recovery from acute episodes.

These burdens often fall disproportionately on women, particularly mothers, reflecting traditional caregiving roles common in South Asian cultures that persist postimmigration. Female caregivers frequently manage both direct patient care and household duties, reinforcing the gendered division of labor described in previous literature [16]. Recognizing this, Hospital-at-Home programs should incorporate gender-sensitive supports, such as education, respite resources, and guidance for safe patient care, to mitigate additional burdens on women.

Another critical finding relates to the health care system education to improve virtual care access. Dahal et al [29] emphasize the role of community organizations in supporting newcomers through education on health care navigation. Participants in this study similarly reported limited awareness of services such as urgent and primary care centers and the 8-1-1 line, resulting in unnecessary reliance on emergency departments. This highlights that health system literacy is a foundational barrier: without addressing these basic knowledge gaps, new service models such as Hospital-at-Home risk reproducing existing inequities rather than improving access.

Chowdhury et al [31] identified concerns about virtual care creating distance between patients and providers, a theme also present in this study. Participants feared that reduced in-person interactions could diminish care quality and trust. This included concerns about developing weaker relationships with providers due to the impersonal nature of virtual interactions, which participants felt could compromise the depth of care and empathy received. This concern reinforces the importance of maintaining strong communication strategies within Hospital-at-Home models to ensure services remain patient-centered and do not further alienate those who experience barriers to care.

While age, gender, and immigration status emerged as individual factors shaping experiences with Hospital-at-Home, our findings underscore the importance of viewing these dimensions through an intersectional lens. For example, older newcomers with limited English proficiency described significant challenges with both language and digital literacy, creating compounded barriers to adoption that were distinct from those faced by younger, Canadian-born participants who reported greater comfort with system navigation and technology use. Similarly, caregiving responsibilities were disproportionately reported by women in multigenerational households, highlighting how gendered expectations intersect with cultural norms and family structures to intensify the caregiving burden. Recent immigrants, regardless of age, also described limited awareness of alternatives to the emergency department, illustrating how immigration status interacts with health system literacy to shape health care–seeking behaviors. These examples illustrate that barriers to Hospital-at-Home adoption are not experienced uniformly across South Asian communities but rather reflect the cumulative and intersecting effects of multiple social identities. Recognizing these layered experiences is critical for designing flexible, equity-oriented models that can adapt to the diverse needs within South Asian populations.

Implications for Hospital-at-Home Implementation

Despite substantial evidence demonstrating the effectiveness of Hospital-at-Home programs in improving patient outcomes and reducing health care costs, significant knowledge gaps remain regarding how to scale these services across diverse populations [32]. This study addresses this gap by exploring the adoption barriers and enablers of Hospital-at-Home services among South Asian communities in the Fraser Health region, offering insights for culturally tailored implementation strategies.

Increasing awareness of Hospital-at-Home services in ways that resonate with South Asian communities is essential to enhance willingness to participate. Findings from this study indicate that South Asian communities prefer learning about Hospital-at-Home through trusted channels such as community organizations, cultural events, and peer testimonials. Additionally, benefits such as increased privacy at home, proximity to family members, and greater control over aligning care with cultural and religious practices were identified as factors that enhance the appeal of Hospital-at-Home services for South Asian patients.

From a systems perspective, raising awareness about how Hospital-at-Home fits within the broader health care system is equally important. Educating patients and caregivers about the structure of health care services and how Hospital-at-Home complements traditional hospital care can build trust, increase understanding of its role in acute care delivery, and foster confidence in available escalation pathways. This system-level approach can help patients see Hospital-at-Home as an integral part of a coordinated healthcare continuum rather than as a stand-alone service.

Addressing individual barriers is also critical for sustainable and equitable adoption. Limited digital literacy skills among South Asian populations can be mitigated through multilingual education and training strategies that accommodate diverse learning styles. Designing user-friendly digital interfaces for Hospital-at-Home services is also critical to ensure accessibility. Furthermore, the study highlights the importance of explicitly considering family members and caregivers in Hospital-at-Home implementation. This includes developing caregiver-focused training materials, addressing family dynamics in care planning, and regularly evaluating caregiver experiences.

This study further emphasizes that preferences and attitudes toward Hospital-at-Home services are not uniform across South Asian communities, nor do they universally apply to all patient needs. For instance, while some participants appreciated the comfort and cultural familiarity of home-based care, others voiced significant concerns around privacy, data security, and the adequacy of clinical care delivered virtually. These diverging views underscore the need for flexible, patient-centered models that allow for individual choice.

Housing environments emerged as an important contextual factor influencing the feasibility of Hospital-at-Home care. Participants described challenges such as stairs, limited space for medical equipment, and crowded or multigenerational households, all of which could constrain the safety and acceptability of receiving acute care at home. These findings point to the need to consider how physical living environments intersect with the delivery of virtual hospital models. While a detailed policy analysis is beyond the scope of this study, it is important to acknowledge that housing affordability and adequacy are pressing issues in Canada, particularly in urban centers such as Surrey. Future research should explore how these broader structural determinants, including housing conditions, shape equitable access to home-based care models and whether supportive housing or community infrastructure can mitigate such barriers.

To ensure equitable and effective implementation, the design of Hospital-at-Home programs must incorporate sufficient resources to continuously engage diverse patient, provider, and community voices. This ongoing feedback process is critical for adapting services to reflect the cultural nuances, varied needs, and evolving preferences of South Asian populations, enabling the development of culturally informed strategies that foster sustained adoption and trust in Hospital-at-Home services. This study's qualitative findings informed a subsequent Experience-Based Co-Design initiative, which actively engaged South Asian community members in the cocreation of culturally responsive solutions to the challenges identified. Specific outcomes of this process included the development of culturally adapted training materials, multilingual communication tools, and revised care workflows to strengthen trust and clarity in care delivery. By linking this foundational research with participatory design methods, the broader project not only addressed knowledge gaps but also demonstrated a model for translational research, showing how community-engaged approaches can directly inform the design of culturally safe virtual care innovations.

Policy and System-Level Considerations

The findings from this study highlight the need for equity-focused policies and system-level interventions to enhance virtual care accessibility for South Asian communities in Fraser Health and beyond. A central implication is that health system literacy must be addressed as a prerequisite to virtual adoption; simply introducing Hospital-at-Home services without parallel education on how and when to use them may not improve access. South Asian patients may face barriers due to limited English proficiency and a lack of access to culturally competent care [33]. Policies should mandate the integration of multilingual virtual health services, including real-time interpretation and culturally responsive health communication strategies. This aligns with Fraser Health's ongoing efforts to improve accessibility through initiatives such as the SAHI.

Another critical policy recommendation is the standardization of digital literacy support within virtual care services. The findings of this study identified the problematic impact of technological barriers, particularly among older adults and individuals unfamiliar with digital tools. Policies should ensure that virtual care platforms are user-friendly and that comprehensive digital literacy training programs are offered as part of virtual hospital enrollment. This could include step-by-step tutorials, one-on-one technical assistance, and hybrid learning options that combine online and in-person training [34]. In addition, although only a minority of participants mentioned it, data privacy and security remain essential considerations in the design of virtual hospital services. Policies should ensure that multilingual and culturally tailored digital platforms also meet high standards of technological ethics, including secure data management and patient confidentiality, to build and maintain trust among diverse communities.

System-level reforms should focus on streamlining care pathways for virtual hospital patients. Participants expressed concerns about delays in care escalation and emergency response times. To address this, policies should define clear transition protocols between virtual and in-person hospital care, ensuring that patients requiring higher levels of medical attention can be seamlessly transferred. Emergency response times can be optimized through virtual triage systems to ensure virtual hospital patients receive timely interventions when needed [35]
.

An additional consideration is the inclusion of family caregivers in virtual health planning. Individuals in South Asian households rely on family members and loved ones for health care navigation and support. Virtual care policies should recognize and integrate the role of caregivers, facilitating access to training, resources, and decision-making tools to support patient care effectively.

Finally, data-driven policy improvements are essential for ensuring that virtual care models address the specific needs of South Asian communities. These findings of this study highlight the importance of ongoing patient engagement and feedback mechanisms. There is a clear need to implement structured monitoring and evaluation frameworks that assess adoption, patient experiences, health outcomes, and engagement levels in virtual care programs. These insights should inform policy refinements to ensure equitable and effective virtual health services.

Strengths and Limitations

This study provides novel insights into Hospital-at-Home acceptability among South Asian communities in Canada, a population underrepresented in Hospital-at-Home research. Semistructured interviews in multiple languages enabled diverse perspectives across gender, age, and immigration status. Importantly, the study included voices of marginalized groups—such as elders, recent immigrants, and family caregivers—who are often excluded from virtual health research.

While this study provides valuable insights into the experiences of South Asian communities with virtual health care, several limitations should be noted. Recruitment challenges limited geographic diversity, with most participants residing in Surrey. As access to health services and digital infrastructure varies across municipalities, future studies should aim for broader regional representation to capture a wider range of experiences.

Another limitation is the reliance on referrals from PICS, which may have introduced selection bias. Participants recruited through PICS are likely to have stronger community connections and greater awareness of health resources, leading to a sample more engaged with the health care system than the broader South Asian population. This may result in an underrepresentation of those facing greater systemic and socioeconomic barriers. Nonetheless, the inclusion of participants with limited English proficiency, lower digital literacy, and newcomer status partially mitigates this concern by ensuring that marginalized perspectives were represented.

The reliance on online interviews is another limitation. While virtual interviews increased accessibility for some, they may have excluded participants uncomfortable with digital technologies or lacking reliable internet access, potentially missing the perspectives of those most affected by digital exclusion. Future research should incorporate alternative data collection methods, such as in-person or telephone interviews, to ensure more inclusive participation.

Despite these limitations, this study offers valuable insights into the structural and cultural barriers influencing virtual health adoption among South Asian communities. Future research should use diverse recruitment strategies and mixed methods approaches to enhance sample representativeness and ensure that equity-driven virtual care models are more inclusive of underserved populations.

Conclusion

This study highlights the complex and context-specific factors that influence South Asian communities' engagement with Hospital-at-Home and virtual care services. Barriers such as limited digital literacy, language challenges, caregiver burden, and lack of awareness about health care alternatives must be addressed through culturally and linguistically responsive strategies. At the same time, key facilitators such as trust in health care providers, the comfort of home-based care, and alignment with cultural and familial values offer important starting points for building more inclusive models of care. Ensuring the equitable implementation of Hospital-at-Home services will require sustained investment in community-specific education, caregiver inclusion, and ongoing feedback mechanisms that reflect the diversity within South Asian populations.

Acknowledgments

The authors sincerely thank the South Asian community members who participated in this study and shared their experiences, as well as the Steering Committee for their guidance and support throughout the project.

Funding

This work was supported by Health Research BC and the Health Research Foundation of Innovative Medicines Canada Team Grant in Virtual Care. The funders had no role in the study design, data collection, analysis, interpretation, writing of the manuscript, or the decision to submit the paper for publication.

Data Availability

The deidentified interview transcripts and other data generated or analyzed during this study are available from the corresponding author on reasonable request, subject to ethical approval and participant confidentiality agreements. Due to the sensitive nature of the data and the small community context, data cannot be publicly shared.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full interview guide.

[DOCX File, 17 KB - humanfactors_v13i1e79675_app1.docx]

Checklist 1
COREQ checklist.
[PDF File, 1171 KB - humanfactors_v13i1e79675_app2.pdf]

References

1. Census profile, 2016 census. Statistics Canada. URL: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E [accessed 2025-04-16]
2. Bainey KR, Jugdutt BI. Increased burden of coronary artery disease in South-Asians living in North America. Need for an aggressive management algorithm. Atherosclerosis 2009 May;204(1):1-10. [doi: 10.1016/j.atherosclerosis.2008.09.023] [Medline: 18980768]
3. Rana A, de Souza RJ, Kandasamy S, Lear SA, Anand SS. Cardiovascular risk among South Asians living in Canada: a systematic review and meta-analysis. CMAJ Open 2014 Jul;2(3):E183-E191. [doi: 10.9778/cmajo.20130064] [Medline: 25295238]
4. Palaniappan L, Garg A, Enas E, et al. South Asian cardiovascular disease & cancer risk: genetics & pathophysiology. J Community Health 2018 Dec;43(6):1100-1114. [doi: 10.1007/s10900-018-0527-8] [Medline: 29948525]
5. South Asian health report. Fraser Health. 2015. URL: https://www.fraserhealth.ca/-/media/Project/FraserHealth/FraserHealth/Health-Topics/South-Asian-Health/201606_South_Asian_Health_Report.pdf [accessed 2025-10-14]
6. Islam F, Khanlou N, Tamim H. South Asian populations in Canada: migration and mental health. BMC Psychiatry 2014 May 26;14:154. [doi: 10.1186/1471-244X-14-154] [Medline: 24884792]
7. Godley J. Everyday discrimination in Canada. Can J Sociol 2018;43(2):111-142. [doi: 10.29173/cjs29346]
8. van Draanen J, Shafique A, Farissi A, et al. How to offer culturally relevant type 2 diabetes screening: lessons learned from the South asian diabetes prevention program. Can J Diabetes 2014 Oct;38(5):329-333. [doi: 10.1016/j.jcjd.2013.11.008] [Medline: 24797496]
9. Jennett PA, Affleck Hall L, Hailey D, et al. The socio-economic impact of telehealth: a systematic review. J Telemed Telecare 2003;9(6):311-320. [doi: 10.1258/135763303771005207] [Medline: 14680514]
10. Melman A, Maher CG, Machado GC. Virtual hospitals: why we need them, how they work and what might come next. J Physiother 2021 Jul;67(3):156-157. [doi: 10.1016/j.jphys.2021.06.018] [Medline: 34148815]
11. Ryan PP, Hawkins KL, Altman S, et al. A novel virtual hospital at home model during the coronavirus disease 2019 (COVID-19) pandemic. Infect Control Hosp Epidemiol 2021 Sep;42(9):1140-1142. [doi: 10.1017/ice.2020.435] [Medline: 32829735]
12. Shepperd S, Doll H, Angus RM, et al. Admission avoidance hospital at home. Cochrane Database Syst Rev 2008 Oct 8(4):CD007491. [doi: 10.1002/14651858.CD007491] [Medline: 18843751]
13. Chua CMS, Ko SQ, Lai YF, Lim YW, Shorey S. Perceptions of hospital-at-home among stakeholders: a meta-synthesis. J Gen Intern Med 2022 Feb;37(3):637-650. [doi: 10.1007/s11606-021-07065-0] [Medline: 34363185]
14. Patel HY, West DJ. Hospital at Home: an evolving model for comprehensive healthcare. Glob J Qual Saf Healthc 2021 Nov;4(4):141-146. [doi: 10.36401/JQSH-21-4] [Medline: 37261225]
15. McIntyre D, Chow CK. Waiting time as an indicator for health services under strain: a narrative review. Inquiry 2020;57:46958020910305. [doi: 10.1177/0046958020910305] [Medline: 32349581]
16. Hyman A, Stacy E, Mohsin H, et al. Barriers and facilitators to accessing digital health tools faced by South Asian Canadians in Surrey, British Columbia: community-based participatory action exploration using photovoice. J Med Internet Res 2022 Jan 13;24(1):e25863. [doi: 10.2196/25863] [Medline: 35023842]
17. Jong M, Mendez I, Jong R. Enhancing access to care in northern rural communities via telehealth. Int J Circumpolar Health 2019;78(2):1554174. [doi: 10.1080/22423982.2018.1554174] [Medline: 31066652]
18. Zibrik L, Khan S, Bangar N, Stacy E, Novak Lauscher H, Ho K. Patient and community centered eHealth: exploring eHealth barriers and facilitators for chronic disease self-management within British Columbia's immigrant Chinese and Punjabi seniors. Health Policy Technol 2015 Dec;4(4):348-356. [doi: 10.1016/j.hlpt.2015.08.002]
19. Makowsky MJ, Jones CA, Davachi S. Prevalence and predictors of health-related internet and digital device use in a sample of South Asian Adults in Edmonton, Alberta, Canada: results from a 2014 community-based survey. JMIR Public Health Surveill 2021 Jan 8;7(1):e20671. [doi: 10.2196/20671] [Medline: 33416506]
20. Health authorities. Government of British Columbia. URL: https://www2.gov.bc.ca/gov/content/health/about-bc-s-health-care-system/partners/health-authorities [accessed 2025-04-07]
21. About Fraser Health. Fraser Health. URL: https://www.fraserhealth.ca/about-us/about-fraser-health [accessed 2025-04-07]
22. Census profile, 2021 census of population. Statisics Canada. 2021. URL: https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E [accessed 2025-04-07]
23. PICS. URL: https://pics.bc.ca/ [accessed 2025-04-24]
24. Virtual health. Fraser Health. URL: https://www.fraserhealth.ca/patients-and-visitors/virtual-health [accessed 2025-04-07]

25. South Asian Health Institute (SAHI). Fraser Health. URL: https://www.fraserhealth.ca/health-topics-a-to-z/south-asian-health/south-asian-health-institute [accessed 2025-04-24]
26. Equity, Diversity and Inclusion (EDI). Fraser Health. URL: https://www.fraserhealth.ca/health-topics-a-to-z/EDI [accessed 2025-04-24]
27. Dewey J. Context and Thought, Volume 12: University of California Publications in Philosophy; 1931.
28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
29. Dahal R, Naidu J, Bajgain BB, et al. Patient-identified solutions to primary care access barriers in Canada: the viewpoints of Nepalese immigrant community members. J Prim Care Community Health 2022;13:21501319221141797. [doi: 10.1177/21501319221141797] [Medline: 36476102]
30. Bhalla M, Boutros H, Meyer SB. Aunties, WhatsApp, and “haldi da doodh”: South Asian communities’ perspectives on improving COVID-19 public health communication in Ontario, Canada. Can J Public Health 2022 Dec;113(Suppl 1):46-53. [doi: 10.17269/s41997-022-00712-x] [Medline: 36449223]
31. Chowdhury D, Tong C, Lopez K, Neiterman E, Stolee P. “When in Rome…”: structural determinants impacting healthcare access, health outcomes, and well-being of South Asian older adults in Ontario using a multilingual qualitative approach. Front Public Health 2024;12:1405851. [doi: 10.3389/fpubh.2024.1405851] [Medline: 39741940]
32. Lai YF, Ko SQ. Time to shift the research agenda for Hospital at Home from effectiveness to implementation. Cochrane Database Syst Rev 2024 Mar 5;3(3):ED000165. [doi: 10.1002/14651858.ED000165] [Medline: 38441225]
33. Budhwani S, Fujioka J, Thomas-Jacques T, et al. Challenges and strategies for promoting health equity in virtual care: findings and policy directions from a scoping review of reviews. J Am Med Inform Assoc 2022 Apr 13;29(5):990-999. [doi: 10.1093/jamia/ocac022] [Medline: 35187571]
34. Donetto S, Pierri P, Tsianakas V, Robert G. Experience-based co-design and healthcare improvement: realizing participatory design in the public sector. Des J 2015 Jun;18(2):227-248. [doi: 10.2752/175630615X14212498964312]
35. Gellert GA, Rasławska-Socha J, Marcjasz N, et al. How virtual triage can improve patient experience and satisfaction: a narrative review and look forward. Telemed Rep 2023;4(1):292-306. [doi: 10.1089/tmr.2023.0037] [Medline: 37817871]

Abbreviations

PICS: Progressive Intercultural Community Services
SAHI: South Asian Health Institute

Edited by A Kushniruk; submitted 25.Jun.2025; peer-reviewed by G Wang, M Nisar; revised version received 23.Sep.2025; accepted 16.Oct.2025; published 05.Jan.2026.

Please cite as:
Wong E, Kandola M, Arora K, Sharda H, Deol R, Jung M, Paquin R, Montenegro M, MacPherson M
Hospital-at-Home for South Asian Communities in British Columbia, Canada: Qualitative Interview Study
JMIR Hum Factors 2026;13:e79675
URL: https://humanfactors.jmir.org/2026/1/e79675
doi:10.2196/79675

Original Paper

Bridging Gaps in Women’s Heart Health: User-Centered Needs Assessment Informed by Patient and Clinician Interviews

Christine Jacob[1], PhD; Sangeetha-Rose Puthanveettil[2], MSc; Patrick Vavken[4,5,3*], MD, PD, MSc, MBA; Emel Kaplan[6*], MD, PhD; Christine S Zuern[6*], MD, Prof Dr

[1]Institute for Information Systems, University of Applied Sciences Northwestern Switzerland (FHNW), Windisch, Aargau, Switzerland
[2]University of Applied Sciences Northwestern Switzerland FHNW, Windisch, Switzerland
[3]Vavken Health Labs, Zurich, Switzerland
[4]ETH, Zurich, Switzerland
[5]University of St. Gallen, St. Gallen, Switzerland
[6]Cardiovascular Research Institute Basel (CRIB) and Department of Cardiology, University Hospital Basel, University of Basel, Basel, Switzerland
[*]these authors contributed equally

Corresponding Author:
Christine Jacob, PhD
Institute for Information Systems
University of Applied Sciences Northwestern Switzerland (FHNW)
Bahnhofstrasse 6
Windisch, Aargau, 5210
Switzerland
Phone: 41 56 202 74 64
Email: christine.k.jacob@gmail.com

Abstract

Background: Women with cardiovascular disease (CVD) remain underserved due to gaps in recognition, diagnosis, and care tailored to sex-specific risks. Digital health tools have the potential to address these inequities, but many fail to reflect the distinct needs of women. In a prior review, we assessed 20 CVD apps and 22 wearables and found that only 25% (5/20) of apps and 40% (9/22) of wearables included any sex-specific content, such as hormone cycle tracking and life-stage considerations related to pregnancy or menopause. These findings confirm that current digital tools largely mirror the gender gaps seen in traditional care.

Objective: This study aimed to define the user requirements for a CVD app designed specifically for women. We sought to explore the unmet needs and challenges faced by female patients and their clinicians that current tools fail to address, and also to identify and prioritize features that would be most valuable and feasible to implement.

Methods: We conducted a qualitative study using semistructured interviews to explore the needs, preferences, and expectations of women living with CVD and their treating clinicians. Guided by the human-centered design framework, this work focused on the “Define” phase. A total of 20 participants in Switzerland were interviewed, including 11 women with CVD, 7 cardiologists, and 2 experts in regulatory and reimbursement. Participants were recruited through purposive sampling, and interviews were conducted online between April and July 2025. Thematic analysis was used to synthesize the data, highlighting design priorities and contextual factors relevant for developing a patient-centered and system-aware digital health tool.

Results: The interviews with women living with CVD and cardiologists confirmed the consistent gaps between existing care pathways and the specific needs of female patients. Both groups highlighted the lack of early symptom recognition, insufficient sex-specific guidance, and limited tools tailored to women’s lived experience. While patients prioritized personalized education, emotional support, and features that address hormonal and life-stage–specific risks, clinicians emphasized clinical use, workload integration, and actionable summaries. Success was defined experientially by patients (eg, empowerment and reduced anxiety), and operationally by clinicians (eg, earlier detection and improved adherence). Willingness to pay was moderate among both groups, with patients favoring simplicity and clinicians emphasizing workflow integration and proven clinical use.

Conclusions: These findings highlight the importance of designing an artificial intelligence–enabled CVD app for women that meaningfully integrates patient empowerment with clinical workflows. A dual-value approach is essential, offering personalized tools that address emotional and lifestyle needs for patients, while supporting clinicians with concise, actionable insights. Early reflections on regulatory and reimbursement considerations suggest that a modular, evidence-based rollout strategy would be key for long-term adoption and scale.

(JMIR Hum Factors 2026;13:e82916) doi:10.2196/82916

KEYWORDS

eHealth; mobile health; mHealth; digital health; user research; technology adoption; technology implementation; qualitative study; usability; cardiovascular disease; women's health; human centered design

Introduction

Background

Cardiovascular disease (CVD) remains the world's leading cause of death, claiming nearly 18 million lives each year and impacting more than half a billion people globally [1-3]. Despite decades of progress in cardiology, these numbers underscore a persistent challenge: our current approaches to prevention, diagnosis, and treatment are not reaching everyone equally [3]. One critical blind spot is the under-recognition of sex-based physiological differences in cardiovascular health [4]. Women, in particular, are often underdiagnosed, undertreated, and underserved, partly due to atypical symptom presentation and a longstanding male-centric model of research and care [4]. This has serious consequences, including delays in diagnosis and poorer outcomes for women across the CVD continuum [5].

Digital health technologies offer a promising avenue to change this narrative. Mobile health (mHealth) tools, remote patient monitoring, wearable sensors, and artificial intelligence (AI)–driven decision support systems are reshaping how individuals manage their cardiovascular risk, bringing prevention, detection, and self-care into everyday life [1]. According to the World Health Organization, mHealth refers broadly to health care and public health services supported by mobile devices such as smartphones, wearables, and wireless sensors [6]. These tools can improve access to timely care, enable more tailored interventions, and potentially reduce long-term costs to the health system [7].

Evidence is steadily accumulating in favor of mHealth for cardiovascular care. For example, a systematic review by Coorey et al [8] found that mobile apps can support better blood pressure (BP) control, encourage healthy dietary habits, and reduce hospital readmissions for patients with CVD. Smartphone-based photoplethysmography, a technology that measures changes in blood volume using infrared light, has shown promise in detecting atrial fibrillation (AF) and assessing heart rate (HR) variability, offering an accessible, noninvasive, and scalable solution for early risk detection [9].

Wearable devices are also gaining ground as tools for continuous cardiovascular monitoring. Positioned on the wrist, chest, or hip, these devices can monitor HR, blood oxygen levels, sleep patterns, and physical activity using either photoplethysmography or ECG technology [10,11]. For instance, a study by Guo et al [12] involving more than 187,000 users identified more than 260,000 potential AF episodes, with confirmatory testing validating the diagnosis in most cases. A broader review of smartwatch-based interventions echoed these findings, highlighting improvements in lifestyle behaviors, medication adherence, AF detection, and reductions in unplanned hospitalizations [13].

Current State of CVD Apps and Wearables

Building on the well-documented unmet needs in CVD for women, we evaluated how effectively existing digital health tools, specifically mobile apps and wearable devices, address sex-specific factors in CVD. To do this, we conducted a structured review of 20 patient-facing CVD apps and 22 commercially available wearables. Each tool was assessed using the foundational and contextual dimensions of the sociotechnical framework for evaluating patient-facing eHealth interventions [14], which emphasizes both technical functionality and integration into real-world health contexts.

The results of this assessment, published in a separate study [15], revealed a significant gap: only 25% (5/20) of the reviewed apps and 40% (9/22) of the wearables incorporated sex-specific content. This included considerations such as the impact of hormonal changes, menopause, or pregnancy on cardiovascular health, factors known to influence symptom presentation, disease progression, and treatment needs in women. These findings reinforce the conclusion that digital health tools are not exempt from the systemic gaps that characterize traditional CVD care pathways [15]. Rather, they mirror the underrepresentation of women's needs in CVD. Addressing this gap is critical if we aim to develop digital interventions that are both equitable and clinically effective for women living with CVD.

Objectives

To address this gap, this study explored the user needs and requirements for a digital health app designed specifically for women with CVD. The primary objectives were (1) to understand the specific challenges and unmet needs that female patients and health care professionals encounter in cardiovascular care that are not adequately addressed by current digital tools, and (2) to identify and prioritize the features and functionalities that a new app should incorporate to better support sex-specific cardiovascular health management. Furthermore, the study did not only focus on the preferences and expectations of the primary users, such as patients and clinicians, but also considered broader system-level enablers and constraints, including clinical integration, regulatory requirements, and reimbursement considerations, to guide a practical and scalable development strategy.

Methods

Overview

We adopted a qualitative research approach to explore the nuanced needs and expectations of both patients and clinicians. This methodology was chosen for its strength in capturing complex, context-dependent experiences and sociocultural factors that are often overlooked by quantitative methods [16].

Qualitative methods are increasingly used in health services and technology research, as they allow researchers to uncover the “why” behind user behaviors and preferences [17]. In our case, this approach helped surface the specific ways in which digital tools can support women across the CVD journey, and how clinicians view their potential integration into care pathways. These insights provide a strong foundation for user-centered feature development and iterative design.

Scope and Conceptual Framework

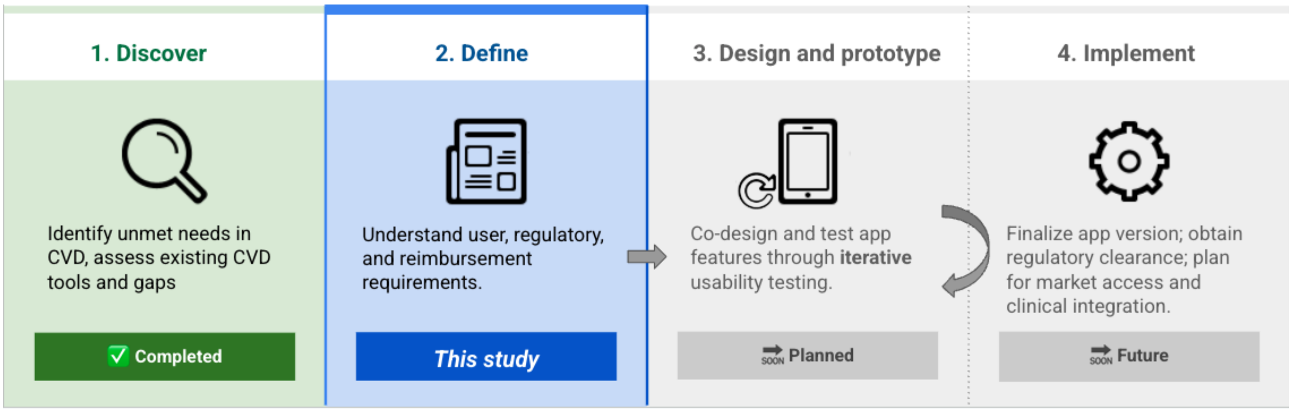
This study was guided by the human-centered design (HCD) framework, which places the needs, experiences, and preferences of end users, here, women living with CVD and the clinicians who care for them, at the core of innovation [18-20]. HCD is a structured, iterative methodology that unfolds across 4 key phases: discover, define, design and prototype, and implement. Each phase plays a critical role in ensuring that health technologies are developed not only for users, but with them [21,22].

The discover phase, which focuses on identifying the problem space and understanding unmet needs, was addressed in our previously published study that assessed the extent to which existing CVD apps and wearables account for sex-specific considerations [15]. That foundational work revealed significant gaps in digital tools for women with CVD, particularly around life-stage–specific guidance and personalized risk tracking [15].

Building on those findings, this study focused on the define phase. This stage aimed to deepen understanding of user needs through qualitative inquiry, synthesize priorities across user groups, and translate insights into clear design criteria. Specifically, we investigated the expectations, preferences, and contextual considerations of both female patients and clinicians to guide the conceptualization of a CVD app tailored to women’s unique health trajectories.

The subsequent design and prototype and implementation phases, where iterative development, user testing, and real-world deployment take place, are beyond the scope of this study but will build directly on the requirements defined in this study. Figure 1 illustrates the HCD process guiding this study, with this study focusing specifically on the define phase, following earlier gap analysis in the discover phase.

Figure 1. Human-centered design approach for developing an app for women with cardiovascular disease. CVD: cardiovascular disease.



Sampling Strategy and Participant Recruitment

We used purposive sampling, a commonly used strategy in qualitative research aimed at capturing rich, experience-based insights from relevant stakeholders [23]. Patient participants were recruited in collaboration with the Women’s Heart Health Program at the University Hospital Basel, a specialized cardiology outpatient program focusing on women's cardiovascular health. Inclusion criteria required participants to be women aged 18 years or older, have a confirmed diagnosis of CVD, access to email and the internet, and be comfortable using a teleconferencing tool.

Eligible patients were approached directly by the clinical staff in the Women’s Heart Health Program, who explained the study and obtained written informed consent. Once consent was given, the signed forms and contact details were forwarded to the core research team at the University of Applied Sciences and Arts Northwestern Switzerland, who subsequently managed the coordination and communication with participants.

Cardiologists were recruited independently by the research team through targeted online searches and professional outreach via email. The participating cardiologists were recruited from 5 different hospitals and clinics, reflecting 3 care settings across Switzerland, including university hospitals, cantonal hospitals, and private practices, offering a diverse clinical perspective. And unlike clinical or patient experience, which can vary widely, regulatory and reimbursement requirements are governed by defined legal frameworks. Therefore, expert input was sought from two highly specialized professionals with national and EU (European Union)-level expertise, who provided focused guidance on certification and financing pathways for digital health tools in Switzerland and Europe.

All participants received study information and consent materials in both English and German and were given the option to choose their preferred language. Interview duration ranged from 28 to 63 minutes, depending on the depth of participant responses. Recruitment took place between April and June 2025 and continued until thematic saturation was reached, that is, when no new themes emerged from additional interviews [23,24]. This was assessed continuously during data collection and analysis. For both patient and clinician interviews, saturation was monitored separately. We noted that by the tenth patient and sixth clinician interview, no substantially new themes were identified, and the final interviews largely confirmed and elaborated on existing patterns. The English version of the participant information sheet and consent form is included in Multimedia Appendix 1.

A total of 20 participants based in Switzerland were recruited: 11 women living with CVD, 7 cardiologists, 1 regulatory expert, and 1 reimbursement expert. The sample characteristics of the participating patients and clinicians are summarized in Multimedia Appendix 2.

Data Collection and Synthesis

Data were collected through in-depth, semistructured interviews conducted online between April and July 2025. In total, 2 tailored interview guides were developed: 1 focused on the patient experience and day-to-day disease management, and 1 directed at clinicians, emphasizing workflow, system integration, and implementation considerations. Both guides are provided in Multimedia Appendix 3.

A structured thematic analysis was conducted following the framework by Braun and Clarke [25,26], including familiarization with the data, initial coding, theme development, review, refinement, and final synthesis (see Multimedia Appendix 4 for detailed steps). We used a hybrid coding approach, combining deductive codes informed by prior work and study objectives, with inductive codes that emerged from the data, as described in Multimedia Appendix 4. Codes were grouped into broader thematic categories and, where relevant, mapped to stages of the patient journey (eg, early diagnosis support, treatment, and long-term management). Coding was supported using NVivo (Lumivero) qualitative data analysis software. Patient and clinician interviews were conducted and primarily analyzed by SP, while CJ and PV led the interviews with the regulatory and reimbursement experts. CJ further refined the thematic analysis, synthesis, and reporting. To mitigate potential researcher bias, coding was triangulated across the 3 coauthors (SRP, CJ, and PV), and discrepancies were resolved through iterative discussion. Any coding discrepancies between SP and CJ were addressed through discussions with PV until a consensus was reached. This process spanned from May to August 2025.

Ethical Considerations

The Ethics Committee of Northwest and Central Switzerland determined that ethics approval was not needed for this study, according to the Federal Act on Research involving Human Beings, article 2, paragraph 1 (reference number Req-2025-00491). All participants were briefed about the research background and signed a consent form agreeing to participate. Participant data were anonymized, access to identifiable source files was restricted to the first and second authors, and no participants received financial compensation, with the exception of the regulatory and market access experts who received consulting fees for their professional advisory contributions.

Results

Understanding the User Journey and Unmet Needs

The patient interviews revealed a consistent pattern of late recognition, emotional burden, and lack of gender-sensitive care in the management of CVD. Many women described experiencing a delayed or unclear diagnosis, often attributed to atypical symptom presentation and a lack of accessible, sex-specific information. Many had to self-educate or navigate their condition alone, especially at the outset, before a formal diagnosis was made. Emotional distress, stress, and fear, especially post diagnosis, were prominent themes. Participants spoke about the emotional toll that followed their diagnosis, including feelings of shock, sadness, fear, and uncertainty about the future. For some, the psychological burden was compounded by life-altering restrictions, such as being advised against future pregnancies or having to give up previously enjoyed activities like vigorous exercise.

There was a widespread perception among the participants that hormonal and life-stage influences (eg, pregnancy and menopause) are underacknowledged in treatment pathways. Only 18% (2/11) of patients received any sex-specific guidance, mostly about the potential dangers of getting pregnant with heart disease. Participating patients stated that they generally struggle with lifestyle adjustments, medication adherence, and managing comorbidities. When discussing the complexity of symptom management, participants frequently described difficulties in understanding medical instructions and uncertainty about how to respond to different symptoms or side effects. This was further complicated by external factors such as stress, an acknowledged trigger that patients knew could worsen their condition, yet felt largely powerless to control.

Digital tools are underused; only 18% (2/11) of patients had experience using CVD apps or wearables, for example, for ECG measurement, while the others expressed confusion, lack of awareness, or anxiety about using them. All participating patients agreed that gender-specific tracking would be beneficial. Table 1 highlights the main challenges and unmet needs faced by women with CVD along the care continuum, their current use of digital health tools, and how many participants referenced each theme in the interviews.

Table 1. Understanding the patient journey and unmet needs (n=11).

Themes and subthemes	Prevalence, n (%)
Early symptoms and diagnosis	
Unfamiliar with the disease at diagnosis	7 (64)
Atypical or no symptoms prior to diagnosis	6 (55)
Self-initiated education (eg, googling and support groups)	5 (45)
Emotional distress after diagnosis (shock, sadness, and fear)	6 (55)
Stress and psychological burden of the disease	5 (45)
CVD[a] perceived as a burden or life-changing event	6 (55)
Life stage and gender-specific care	
Not informed or addressed by HCPs[b]	9 (82)
Was informed about the impact of life stages (eg, pregnancy and menopause)	2 (18)
Perceived treatment as male-centered	4 (36)
Belief that women’s CVD symptoms are often dismissed	3 (27)
Ongoing disease management challenges	
Difficulty with lifestyle and diet changes	6 (55)
Struggles with medication adherence or side effects	5 (55)
Difficulty understanding symptoms and instructions	4 (36)
Difficulty managing comorbidities	3 (27)
Anxiety before appointments because of uncertainties (eg, on potential progression)	2 (18)
Dealing with constant emotional overwhelm and fatigue	3 (27)
Use of existing digital tools	
Has not used any apps for CVD	9 (82)
Not aware of any helpful CVD apps	7 (64)
Prefers face-to-face care over digital tools	4 (36)
Feels overwhelmed or anxious by the idea of apps	3 (27)
Has used a step counter, hydration reminder, or menstrual tracker	3 (27)
Has used a CVD-related app or wearable (eg, for ECG[c] measurement)	2 (18)

[a] CVD: cardiovascular disease.

[b] HCP: health care professional.

[c] ECG: electrocardiogram.

Clinicians reported several diagnostic challenges stemming from atypical symptom presentation in women, compounded by limited awareness and insufficient referral pathways. They emphasized the mismatch between traditional risk models and real-world female presentations, which are often underrecognized or misattributed. Time constraints and health care system overload further exacerbate diagnostic delays. Despite acknowledging the lack of tools to capture hormonal or life-stage influences, 5 of 7 (71%) participating clinicians reported that they continue to follow standard monitoring protocols without sex-specific adjustments.

Only 1 of 7 (17%) interviewed cardiologists recommended CVD apps, while 6 of 7 (86%) cited lack of familiarity or trust in their use, with usability for older patients as the main concern. Integration of digital tools into clinical workflows remains controversial; some see potential for structured summaries and AI-assisted risk insights, while others worry about time burden or data security. There was strong agreement on the need for decision support tools that surface earlier warnings during high-risk windows such as menopause, pregnancy, or postevent follow-ups, with 5 of 7 (71%) participants citing this as an existing gap. Table 2 presents the key challenges and unmet needs encountered by clinicians throughout the clinical workflow, alongside their current engagement with digital health tools, and how many participants referenced each theme in the interviews.

Table 2. Understanding the clinician workflow and unmet needs (n=7).

Themes and subthemes	Prevalence, n (%)
Diagnostic challenges	
Atypical symptoms in women make the diagnosis more challenging	7 (100)
Lack of awareness and, accordingly, referrals	5 (71)
Limited time and capacity in care	3 (43)
Sex-specific gaps	
Symptoms misattributed or dismissed	7 (100)
Treatment paths not sex-sensitive and mostly male-focused	5 (71)
Monitoring practices	
Standardized tests, no sex differences	5 (71)
Limited tools for life-stage tracking in women (eg, pregnancy and menopause)	3 (43)
Workflow integration needs	
Desire for structured reports	6 (86)
Need decision support during diagnosis and follow-ups	5 (71)
EHR[a] integration	3 (43)
Use of existing digital tools	
Does not recommend CVD[b] apps to their patients	6 (86)
Low awareness of validated apps	6 (86)
Too many apps with low usability for older adults	5 (71)

[a] EHR: electronic health record.

[b] CVD: cardiovascular disease.

Desired Features and Functionality

Based on the patient interviews, several key themes emerged regarding desired features and functionalities for a CVD app tailored to women. Patients emphasized the importance of integrated support across the full care continuum, from early symptom awareness and diagnosis to daily management and communication with health care providers.

Core priorities included the ability to track vital signs, log symptoms, and receive tailored educational content specific to women's cardiovascular health. Medication reminders, stress management features, and dietary guidance were commonly requested. Also, 9 of 11 (82%) patients expressed interest in syncing the app with wearables to streamline data collection and support longitudinal tracking. Communication features, such as automated report generation and the ability to share real-time data with health care providers, were considered highly valuable, with 8 of 11 (73%) patients supporting real-time feedback or alerts when symptoms are concerning. Importantly, all participants favored sex-specific, personalized recommendations, but some voiced concerns about privacy, complexity, and the risk of being overwhelmed by notifications. Table 3 summarizes the key features and functionalities patients desire in a digital health solution, along with the main barriers and concerns, and indicates how many participants raised each point during the interviews.

Based on the interviews with cardiologists, several priorities emerged. The overarching emphasis was on improving diagnostic precision, treatment adherence, and patient-provider communication, while minimizing time burden and clinical noise. Clinicians expressed strong interest in receiving structured patient-generated data (eg, BP, symptom trends, medication adherence, and hormonal cycle data), especially when aggregated into concise, longitudinal reports. There was broad consensus (7/7, 100%) on the value of sex-specific insights (eg, menopause and pregnancy risks), particularly if tailored to life stages and actionable. However, real-time monitoring or alerts received more mixed responses. Most clinicians (5/7, 71%) preferred periodic summaries over continuous alerts, citing concerns around workload, liability, and alert fatigue.

Communication through the app was not favored by most clinicians, who instead preferred to retain current channels such as email, phone, or in-person visits. While 6 out of 7 (86%) found the ability to customize patient goals important, they highlighted that this must not increase cognitive or administrative burden. Concerns were raised about data quality, patient over-reliance on technology, and integration with existing clinical infrastructure (eg, electronic health records [EHRs]). Despite some openness to features such as AI-driven alerts or remote monitoring, most clinicians emphasized the need for evidence of clinical use and a clear focus on reducing, not increasing, workload. Table 4 summarizes the key features and functionalities clinicians desire in a digital health solution, along with the main barriers and concerns, and indicates how many participants raised each point during the interviews.

Table 3. Desired features and functionality: the patient’s perspective (n=11).

Themes and subthemes	Prevalence, n (%)
Diagnosis support and education	
Educational content tailored to women's heart health	10 (91)
Summaries and videos explaining medications or symptoms	3 (27)
Symptom awareness and early management	
Symptom logging	3 (27)
Hormonal cycle insights	2 (18)
Personalized alerts based on symptoms or wearable input	5 (45)
Treatment management and adherence support	
Medication reminders	7 (64)
Diet and lifestyle guidance	5 (45)
Personalized health recommendations based on lifestyle and symptoms	5 (45)
Appointment reminders	4 (36)
Monitoring and tracking	
Vital signs tracking (BP[a], HR[b], ECG[c], etc)	9 (82)
Integration with wearables	9 (82)
Sleep and stress tracking	3 (27)
Cycle tracking	2 (18)
Communication and feedback loops	
Real-time feedback or alerts when symptoms are concerning	8 (73)
Automated report generation	6 (55)
Ability to share data with HCPs[d]	6 (55)
In-app messaging with the care team	4 (36)
Motivation and well-being	
Exercise and mindfulness tips	2 (18)
Community and patient support forums	2 (18)
Barriers and concerns	
App complexity	4 (36)
Privacy concerns	3 (27)
Over-reliance or tech-induced anxiety	3 (27)
Over-generalization of content	2 (18)
No concerns	4 (36)

[a] BP: blood pressure.

[b] HR: heart rate.

[c] ECG: electrocardiogram.

[d] HCP: health care professional.

Table 4. Desired features and functionality: the clinician's perspective (n=7).

Themes and subthemes	Prevalence, n (%)
Treatment planning and customization	
Use of sex-specific guidance (pregnancy and menopause)	6 (86)
Important to customize goals and treatment plans	6 (86)
Monitoring preferences	
Prefer active patient monitoring	3 (43)
Prefer alerts-only model	3 (43)
Would not use the app for monitoring (education use only)	1 (14)
Communication features	
Prefer automated patient reports	5 (71)
In-app direct communication not preferred (prefer phone, email, and in-person follow-up)	5 (71)
Alert preferences	
Prefer periodic summaries only	5 (71)
Prefer both real-time alerts plus summaries	1 (14)
Prefer alerts only for high-risk events	1 (14)
Patient-generated data priorities	
Blood pressure readings	7 (100)
Heart rate and rhythm	5 (71)
Symptom tracking	4 (57)
Medication adherence reports	4 (57)
Hormonal cycle fluctuations	4 (57)
Stress, sleep, and mental health indicators	3 (43)
Exercise and activity levels	3 (43)
Weight and BMI changes	2 (29)
Lifestyle factors (eg, diet and smoking)	2 (29)
CVD[a] risk score integration	2 (29)
Diagnostic data (eg, lab values and cholesterol)	2 (29)
Barriers and concerns	
Excess workload and nonactionable data (noise)	4 (57)
Data accuracy and false alarms	3 (43)
Legal liability if alerts go unaddressed	3 (43)
Poor EHR[b] integration	3 (43)
Privacy and data security concerns	3 (43)
Lack of clinical use	2 (29)
Patients' over-reliance on the app	2 (29)
Motivation and sustained engagement by patients	1 (14)

[a] CVD: cardiovascular disease.

[b] EHR: electronic health record.

Contrasting the two perspectives revealed strong alignment between patients and clinicians on the value of tracking vital signs, providing sex-specific guidance, and offering educational content tailored to women's heart health. However, notable divergences emerged around communication, monitoring preferences, and the diagnostic support role of the app. Patients favored real-time feedback, interactive features, and personalized support, while clinicians expressed concerns about workload, data reliability, and legal liability, preferring structured summaries and limited app-mediated communication. These differences highlight the need for a dual-pathway design that balances patient empowerment with clinical workflow and safety.

Table 5 compares the perspectives of participating women with CVD versus cardiologists regarding desired app features across the patient journey. It highlights areas of alignment and divergence across different phases of the patient journey, and identifies where design focus is most needed to address gaps and support user-centered implementation.

Table 5. User requirements comparison across the cardiovascular disease patient journey (patients vs clinicians).

Feature cluster	Patient perspective	Clinician perspective	Alignment
Symptom onset and early risk tracking	Many patients want tools to log symptoms, track hormonal cycles, and receive alerts for serious changes.	Clinicians track symptoms but prefer periodic summaries over real-time alerts due to workload and liability concerns.	Partial
Diagnosis support	Patients often seek symptom explanations and hope for diagnostic guidance from the app.	Clinicians are wary of misdiagnosis; some see the app as educational but not diagnostic.	Divergent
Treatment adherence and medication	High interest in medication reminders, side effect information, and adherence support.	Clinicians value adherence tracking and side effect reporting if actionable, summarized, and integrated into their workflow.	Strong
Education and empowerment	Nearly all patients want tailored, female-focused educational content (eg, videos, articles, and lifestyle advice).	Clinicians agree that education is a key benefit of the app, especially if it supports better patient engagement.	Strong
Sex-specific and life stage guidance	All patients emphasized the need for advice linked to hormonal changes, pregnancy, and menopause.	Most clinicians strongly support sex-specific features to address overlooked risks and life-stage changes.	Strong
Vital signs and lifestyle monitoring	Most want to track vitals (BP[a] and HR[b]) and sync with wearables for exercise, diet, and sleep insights.	Clinicians prioritize BP and HR and accept wearable data cautiously, questioning accuracy and clinical validity.	Moderate
Personalization and motivation	Strong interest in personalized advice and goal-setting features to support self-management.	Clinicians support customizable goals but are concerned about overburdening users or implying unsupported precision.	Partial
Communication with care teams	Many patients want in-app messaging, shared data, and reminders to feel more connected and supported.	Clinicians prefer structured reports and oppose app-based messaging due to time limits and workload.	Divergent
Monitoring preferences	Most prefer real-time feedback if needed, while some fear over-monitoring or anxiety from alerts.	The majority prefer summary reports; only a few support real-time data access or alerts, citing resource constraints.	Divergent
Trust, privacy, and usability	Simplicity and ease of use matter most; some express concerns about privacy and overalerting.	Clinicians worry about liability, data overload, and integration challenges; they prefer actionable, low-burden tools.	Moderate

[a] BP: blood pressure.

[b] HR: heart rate.

Success Metrics and Willingness to Pay

Patients defined success with the app in terms of greater control, improved understanding, and emotional reassurance. Key success indicators include feeling more in control of their heart health (7/11, 64%), increased awareness of sex-specific symptoms and conditions (6/11, 55%), and reduced anxiety or stress (5/11, 45%), along with better adherence to medication and more personalized guidance. For many, success is also tied to the app's ability to improve communication with health care providers and deliver trustworthy, up-to-date content.

Regarding willingness to pay, the average score was 7.1 on a scale of 1 to 10, with individual responses ranging from 3 to 10. Preferences for pricing models were mixed, with a slight preference for one-time payments (6/11, 55%), reflecting a desire for financial simplicity and predictability. Monthly models were preferred by 4 of 11 (36%) respondents for their flexibility, though some requested a free trial period as a prerequisite.

The primary factors influencing willingness to pay included the app's demonstrated usefulness, degree of personalization, sex- and age-specific features, trustworthiness, and perceived value for improving cardiovascular health. A few respondents flagged affordability and usability as potential barriers, highlighting the importance of designing an accessible and demonstrably beneficial tool. Table 6 provides an overview of patient-reported success metrics and expectations around payment for digital health solutions, including the number of participants who mentioned each aspect during the interviews.

Table 6. Patient-reported success metrics and payment expectations (n=11).

Themes and subthemes	Prevalence, n (%)
Perceived success indicators	
Feeling more in control of heart health	7 (64)
Increased awareness of sex-specific symptoms and conditions	6 (55)
More personalized recommendations	5 (45)
Reduced stress or anxiety	5 (45)
Increased motivation to manage health	4 (36)
Access to reliable and up-to-date health information	3 (27)
Improved medication adherence	2 (18)
Improved communication with doctors (eg, being taken seriously)	2 (18)
Fewer emergency visits	1 (9)
Preferred payment model	
One-time payment	6 (55)
Monthly subscription	4 (36)
Mixed (sees pros and cons in both)	1 (9)
Payment decision drivers	
Proven usefulness and effectiveness	6 (55)
Personalized recommendations or tracking	5 (45)
Availability of sex- or age-specific features	4 (36)
Free trial option	3 (27)
Affordability or current financial situation	3 (27)
Recommendation by a doctor or a trusted source	2 (18)
Simplicity and ease of use	2 (18)
Positive user reviews	1 (9)
Direct communication or data sharing with HCP[a]	1 (9)

[a] HCP: health care professional.

Clinicians defined the success of a CVD app through its impact on patient engagement, adherence, and health outcomes, rather than direct clinical decision-making support. The most frequently cited indicators of success included improved patient engagement and participation in care management (6/7, 86%), greater adherence to treatment (5/7, 71%), and better patient education and symptom recognition (5/7, 71%), particularly regarding sex-specific cardiovascular risks. Several clinicians emphasized the need for such outcomes to be validated through research before adoption in routine practice.

Adoption incentives included features that support remote monitoring, patient education, and AI-supported risk assessments, provided they are accurate and well-integrated. Seamless integration into clinical workflows and EHRs, a simple interface, and support for research and data sharing were also valued.

Regarding willingness to pay, clinicians expressed moderate interest, with a mean score of 5.3 (SD 1.3) on a scale of 1 to 10. Preferences leaned toward monthly or yearly pricing models, with a desire for a free trial period. Willingness to pay was highly conditional on the app's proven benefit to patient outcomes, usability, compliance with data protection, and ability to support research or institutional adoption. Table 7 provides an overview of clinician-reported success metrics and expectations around payment for digital health solutions, including the number of participants who mentioned each aspect during the interviews.

Table 7. Clinician-reported success metrics and payment expectations (n=7).

Themes and subthemes	Prevalence, n (%)
Perceived success indicators	
Improved patient engagement and participation in care management	6 (86)
Improved adherence to treatment	5 (71)
Better symptom recognition and earlier referral	5 (71)
Better patient education and self-management	5 (71)
Reduction in preventable cardiovascular events or hospitalizations	3 (43)
Must be tested in a research context	3 (43)
Easier or faster appointments due to prefilled questionnaires	2 (29)
Long-term lifestyle improvements	1 (14)
Increased patient satisfaction and trust	1 (14)
Must-have features for adoption	
Remote monitoring or real-time patient tracking	4 (57)
Educational content (videos, disease explanations, etc.)	4 (57)
AI[a]-driven risk assessments	3 (43)
Ability to use data for research	3 (43)
Patient engagement tools (behavior change and motivation)	3 (43)
Simple, intuitive user interface	3 (43)
Integration with EHR[b] systems	3 (43)
Preferred payment model	
Monthly payment	3 (43)
Yearly payment (with trial preferred)	2 (29)
One-time payment	1 (14)
No preference	1 (14)
Payment decision drivers	
Proven improvement in health outcomes and adherence	4 (57)
Ability to support research (questionnaires and data export)	3 (43)
Data security and privacy guarantees	2 (29)
Institutional or guideline endorsement	2 (29)
Direct benefit to patients (motivation and education)	2 (29)
Ease of use (for both patients and providers)	2 (29)
Broad adoption or market share	1 (14)
Gender-specific features are fully integrated	1 (14)

[a] AI: artificial intelligence.

[b] EHR: electronic health record.

Regulatory and Reimbursement Considerations

To complement the user requirements gathered from patients and clinicians, we conducted additional interviews with one regulatory expert (SN [nonauthor]) and one reimbursement expert (MF [nonauthor]), both with in-depth knowledge of the Swiss and broader European health care landscapes. This supplementary perspective is critical, as the app will be developed primarily for use in Switzerland, with a medium-term vision for deployment in other European contexts.

The regulatory expert assessed that an AI-enabled CVD app (AI-driven alerts and AI-supported risk assessments), designed to support women across the spectrum of symptom recognition, diagnosis, treatment adherence, and self-management, is likely to be considered a medical device under Swiss law. Specifically, if the app includes features such as AI-driven risk alerts, symptom tracking, and clinical report generation intended to support medical decision-making or influence health outcomes, it would fall under the Medical Device Ordinance as governed by Swissmedic. Depending on the final functionality, the app would likely be classified as a Class IIa medical device, requiring compliance with regulatory obligations related to safety, performance, and quality assurance.

Given that the app would likely use algorithms to generate health risk scores and may eventually integrate with electronic health records or wearable data for clinical review, transparency of its AI functionalities (eg, AI-driven alerts and AI-supported risk assessments) and proper documentation of its performance are essential. Although Switzerland is not part of the EU, it has harmonized its medical device regulatory system with the EU Medical Device Regulation (MDR). Therefore, developers must ensure conformity with Annex I of the MDR, particularly concerning software validation, cybersecurity, and human oversight of AI outputs. Early consultation with Swissmedic, including use of their pre-submission guidance services, was strongly recommended by the expert to confirm classification and identify the most appropriate pathway to conformity assessment.

From a data protection standpoint, the expert added that the app must comply with the Swiss Federal Act on Data Protection, which aligns closely with the EU's General Data Protection Regulation. This includes obtaining explicit consent for processing health data, ensuring transparency in AI-based recommendations, and conducting a Data Protection Impact Assessment if sensitive data are used for profiling or personalized recommendations. Data must be stored securely in accordance with Swiss and European data security standards.

In terms of market access within Switzerland, the reimbursement expert explained that there is currently no established reimbursement pathway for digital health tools of this nature under the standard TARMED system (soon to be replaced by the new tariff system called TARDOC). However, several viable strategies exist. These include partnerships with supplemental insurance providers, integration into occupational health programs, or collaborative pilots with cantonal public health authorities focused on prevention and chronic disease management. Additionally, public-private partnerships involving academic hospitals could support real-world evidence generation, which will be critical for clinical validation and acceptance.

As the app evolves, modular feature deployment (eg, launching initially with educational content and lifestyle tracking before integrating AI-based alerts or EHR connectivity) may offer a lower-risk route to initial uptake. Positioning the app as an adjunct to care rather than a diagnostic tool can help mitigate medico-legal risks and facilitate clinician adoption. For future scaling into EU markets, preparation for CE (Conformité Européenne) marking and consideration of fast-track pathways like Germany's Digitale Gesundheitsanwendungen (Digital Health Applications) process would be key strategic steps. A summary of regulatory and reimbursement considerations is provided in Multimedia Appendix 5, focusing primarily on Switzerland, with additional insights relevant to potential midterm expansion into the EU market.

Discussion

Primary Findings

This study offers detailed insights into how digital health tools can be designed to better serve women with CVD by directly addressing long-standing gaps in diagnosis, treatment, and self-management. What distinguishes this work is its grounding in sex-specific lived experiences and clinician workflows; most existing tools and studies either generalize across populations or fail to capture the unique challenges women face across the cardiovascular care continuum.

Our findings confirmed the diagnostic blind spots long documented in the literature but go a step further by detailing how these gaps are experienced by women, especially the emotional toll of feeling dismissed or misdiagnosed due to atypical symptom presentation, and how clinicians themselves acknowledge the limitations of current care pathways and tools. We also identified a mismatch between what patients expect from digital tools (eg, personalized, real-time, and educational support) and what clinicians view as feasible or desirable (eg, structured summaries and low-disruption integration), echoing broader implementation challenges in digital health but offering concrete, user-validated features to bridge this divide.

Furthermore, this work goes beyond user needs to surface actionable priorities for design, regulation, and reimbursement. By mapping unmet needs to specific app features and clarifying stakeholder requirements, we contribute an applied framework for the development of a digital tool that is not only patient-centered but also clinically and systemically grounded.

Addressing Unmet Needs Across the Care Continuum

Our interviews with patients and clinicians confirmed the consistent disconnect between current cardiovascular care pathways and the lived experiences and needs of women with CVD. Both groups emphasized a clear gap in early recognition and diagnosis of CVD in women, often driven by atypical symptom presentation, a finding well-supported in the literature [4]. Women frequently report vague or nonspecific symptoms (eg, fatigue, anxiety, and back pain) that are either misattributed or dismissed, leading to delayed diagnosis and poorer outcomes. This aligns with clinical studies showing that atypical symptoms in women are associated with underdiagnosis and delayed care, a pattern historically described as the Yentl syndrome and reaffirmed in recent studies [27].

Clinicians similarly confirmed that CVD does not present in women as it does in men, often requiring multiple diagnostic steps and greater clinical vigilance. Yet, most participating clinicians admit to using standardized, male-centered diagnostic tools, with little to no adaptation for sex- or life-stage-specific risks (eg, menopause and pregnancy-related complications). This is consistent with research papers that highlight the lack of integration of sex-specific risk factors into clinical practice and guidelines [28,29].

Our findings also showed that there is a shared perception that awareness and education are insufficient, both for patients and providers. Patients often feel overwhelmed and underinformed, particularly about how hormonal changes might affect their heart health. Clinicians acknowledge this gap, noting that many are not trained to assess or communicate about sex-specific CVD risks, confirming concerns raised in the literature about persistent gender bias in cardiovascular medicine [30].

On the digital health front, while most patients have not used cardiovascular apps, citing lack of awareness, complexity, or emotional overwhelm, many express openness to tools that are personalized, educational, and easy to use. Importantly, all patients agreed that features like hormone-aware symptom tracking and cycle or menopause-specific guidance would be valuable. Clinicians were generally supportive of a digital tool, provided it reduces workload and offers actionable insights. However, many remain skeptical due to workflow integration concerns, unclear clinical use, and insufficient validation, echoing findings from other studies that cite similar adoption barriers [14,31-33].

This shows a strong concordance between user-reported unmet needs in our research and evidence in the scientific literature. The development of a CVD app specifically designed for women presents an opportunity to bridge the identified diagnostic support and management gaps by offering sex-specific decision support and longitudinal symptom tracking across life stages, in a way that integrates meaningfully into both patient journeys and clinician workflows.

Balancing Patient-Centered Design With Clinical Workflow Integration

The synthesis of clinician and patient requirements revealed meaningful overlap in priorities for a CVD app tailored to women, particularly around core features such as BP, HR, cycle and symptom tracking, medication adherence tools, and educational content specific to women's cardiovascular risks. Both groups also expressed interest in incorporating wearable data, lifestyle tracking, and AI-driven alerts, but with key differences in expectations around data flow and engagement. To ensure alignment with human factors (HF) principles, app features and interface components should be conceptualized to minimize cognitive load, support workflow compatibility, and accommodate differentiated engagement preferences between patients and clinicians.

Patients consistently emphasized the importance of holistic, personalized support, especially around lifestyle guidance (eg, diet and stress management), menstrual and hormonal tracking, and emotional well-being. Our recent research, which reviewed 20 commercially available CVD apps, showed that only 25% (5/20) offered sex-specific content, reinforcing that sex-specific app features in CVD are rare and needed [15]. Patients also valued real-time feedback and communication, but with flexibility to self-regulate interaction frequency, a finding supported by previous research [34].

Clinicians, by contrast, largely favored periodic summaries over real-time alerts, citing concerns about information overload, workflow disruption, and medico-legal liability, concerns echoed across several reviews on eHealth implementation challenges [14,32,33,35], underscoring the importance of minimizing cognitive workload and supporting efficient information processing in clinical settings. There was strong support for automated, concise reporting tools to enable faster clinical decision-making, particularly if data is aggregated and actionable (eg, medication nonadherence flags, high BP trends, and cycle fluctuation). Importantly, clinicians stressed the need for sex-specific clinical insights (eg, pregnancy-safe medication guidance), aligning with calls in the literature to address gender gaps in cardiovascular risk assessment and care pathways [36].

Despite agreement on many core features, divergence between patients' and clinicians' requirements remains. Clinicians were skeptical of communication via the app, preferring existing channels (email and phone), whereas many patients desired low-barrier communication tools or asynchronous Q&A features. Additionally, clinician trust in wearable data remains limited due to concerns about accuracy, despite evidence that wearables can support prevention strategies and risk monitoring when used appropriately [10,37].

These findings underscore the need to balance patient-centered design with clinical workflow integration, ensuring that features valued by patients (eg, education, personalization, symptom explanation, and risk prevention) do not impose undue burden on clinicians. The app should prioritize structured reporting, passive data aggregation, and modular engagement options, and be framed as an adjunct, not a replacement, for in-person care. Differentiated user models, with patient-facing tools focused on behavior change, motivation, and real-time support, and clinician-facing tools optimized for decision efficiency and minimal disruption, can support better task-technology fit for both groups. Interface design should adhere to HF principles such as customizability by role, transparency of AI outputs, clear visual hierarchies, and user autonomy in alert settings, all of which were strongly echoed in user preferences. Usability must remain a central design principle, interpreted not only as ease of use but also through HF constructs such as interaction efficiency, information clarity, and minimization of user burden. Successful uptake will hinge on transparent validation, privacy safeguards, and clinical evidence.

Articulating Meaningful Success Narratives

The analysis of success metrics and willingness to pay highlighted distinct yet interrelated priorities between patients and clinicians that carry significant implications for the app's value proposition and sustainable adoption. While both groups identify improved patient adherence, better education, and enhanced engagement as key indicators of effectiveness, they differ in how success is defined and operationalized.

Patients tend to frame success in terms of personal empowerment and perceived well-being, such as increased awareness of sex-specific symptoms and conditions, feeling in control of their heart health, reduced anxiety, and improved ability to interpret symptoms and make informed decisions. Most patients cited greater motivation and personalized insights as important outcomes, and several emphasized a desire for the app to help them feel taken seriously in clinical encounters. These findings align with research demonstrating that digital interventions can significantly boost patient empowerment, knowledge, and self-management in chronic illness [13,38]. This experiential framing of success suggests that patients value both functional benefits (eg, symptom recognition and adherence reminders) and emotional reassurance, and are likely to judge effectiveness through lived experience rather than clinical endpoints alone.

Clinicians, in contrast, emphasized evidence-based outcomes, such as reduced hospitalizations, earlier detection of complications, and measurable improvements in disease management. While some acknowledged softer benefits like improved communication or decreased workload due to better-informed patients, there was a consistent call for formal evaluation of impact, preferably through research trials, factors frequently noted in clinician acceptance studies [32,33]. This divergence highlights the importance of building a dual feedback system, one capturing patient-reported outcomes and engagement metrics, and another tracking clinical indicators that can be aggregated and validated over time.

In terms of willingness to pay, patients showed relatively high interest. Key influencing factors included the app's ability to improve cardiovascular health, offer sex-specific and personalized recommendations, and provide reliable, easily digestible information, echoing user preferences reported in patients' mHealth adoption literature [34]. Several patients noted the appeal of free trial options and transparent pricing, especially in light of economic constraints.

By contrast, clinicians expressed moderate willingness to pay and viewed app adoption as contingent on demonstrated clinical use, alignment with existing systems, and endorsement by institutions or guidelines. Most clinicians would expect the app to be provided at the health system or institutional level, particularly if it supports research or integrates with existing systems, factors shown as crucial for sustained clinician adoption in previous research [32,33,35].

Payment feasibility was primarily seen through the lens of institutional procurement rather than individual expenditure. This indicates that a tiered strategy could be considered, potentially offering direct-to-patient subscriptions for patients and licensing options or integration pathways for clinical and research settings. Such a dual-strategy may combine offering core low-risk features and premium add-ons directly to motivated patients, with an institutional offering that includes more advanced research data modules such as an AI-predictive model, EHR integration, and structured reporting and smart alerts. Free trials, outcome-based pricing, or inclusion in reimbursement schemes could improve uptake across both segments, as emphasized in similar research [39,40].

Overall, the findings suggest that to drive adoption and value across stakeholder groups, the app must articulate distinct success narratives, one rooted in empowerment and lived experience for patients, and one grounded in clinical efficiency and measurable outcomes for providers.

Regulatory and Reimbursement Implications

Incorporating regulatory and reimbursement expertise early in the development process helps ensure that the solution is not only clinically and user-relevant but also aligned with the legal, safety, and financial frameworks required for real-world adoption. Failure to consider these aspects at an early stage often results in substantial implementation barriers and limits the scalability and sustainability of digital health innovations [14,32-35,41].

From a regulatory standpoint, the expert confirmed that if the app includes AI-driven functionalities, such as personalized risk alerts, clinical report generation, or features that support diagnosis or influence medical decision-making, it would likely qualify as a Class IIa medical device under Switzerland's Medical Device Ordinance, harmonized with the EU MDR. This classification implies that the app must meet formal requirements related to software performance, cybersecurity, human oversight, and clinical evaluation, as outlined in MDR Annex I. Given the inclusion of health data processing and algorithmic personalization, the app must also comply with the Swiss Federal Act on Data Protection, which closely mirrors the general data protection regulation, including requirements for explicit user consent and transparency of algorithmic outputs. These recommendations align with recent regulatory reviews emphasizing the increased scrutiny of AI-based medical software and the importance of early consultation with regulatory bodies to validate classification and compliance pathways [42].

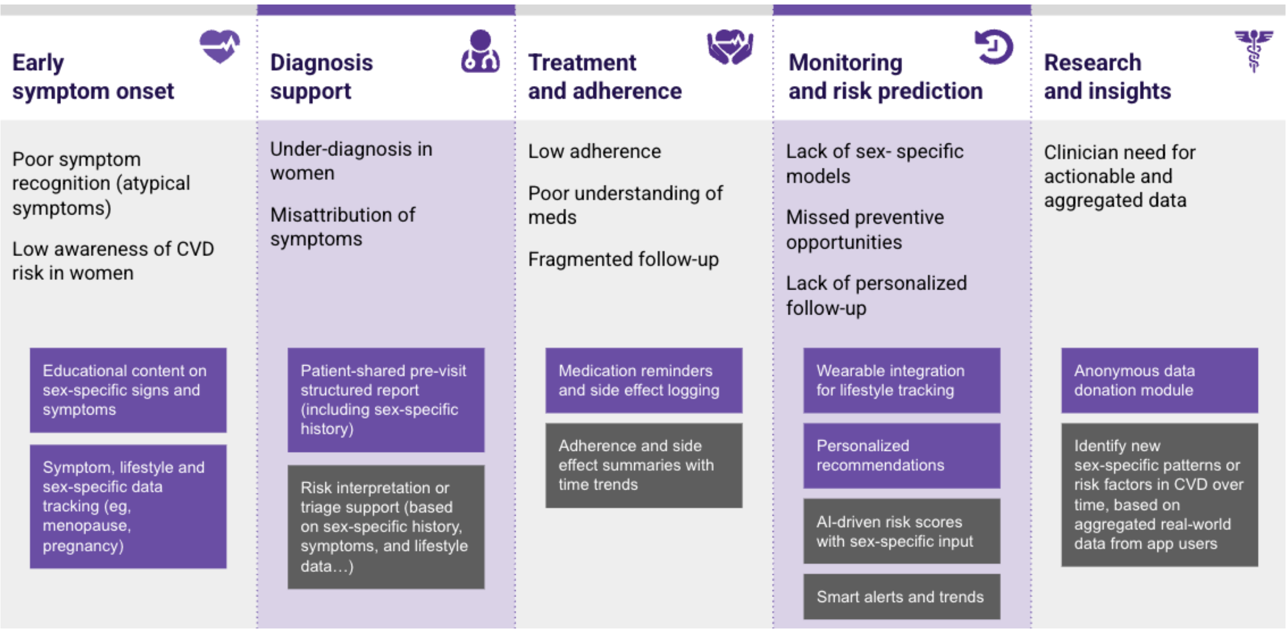
On the reimbursement side, the expert highlighted that Switzerland currently lacks a dedicated reimbursement pathway for digital health applications within the standard outpatient tariff system (TARMED), which is soon to be replaced by the new TARDOC system. In the midterm, TARDOC will be the key framework to monitor. According to the factsheet on digital health applications provided by the Federal Office of Public Health, digital tools, including AI applications, are expected to be classified as part of the infrastructure or personnel services [43]. This means they could potentially be reimbursed under specific TARDOC positions, such as those related to telemedicine. However, several alternative access models were identified, including partnerships with private health insurers, occupational health initiatives, and cantonal prevention programs. These strategies reflect recent calls in the literature for shifts in European digital health financing, which recognize the need for novel evidence-generation mechanisms to support the adoption and scaling of digital tools [44,45]. In Switzerland, reimbursement will likely depend on demonstrated value in real-world settings, particularly in enhancing adherence, promoting prevention, and reducing unnecessary health care use.

Design Implications and Next Steps

To translate the diverse and specific user requirements identified in this study into actionable design features, we mapped patient and clinician needs along the cardiovascular care continuum. Thematic synthesis of interview data revealed critical gaps and opportunities at multiple points in the patient journey, from early symptom recognition to diagnosis support, treatment adherence, ongoing self-management, and broader contributions to cardiovascular research. These findings aligned with persistent shortcomings in the literature, particularly around under-recognition of sex-specific symptoms [46], insufficient tailoring of risk assessment tools, and limited integration of digital health tools into routine care for women with CVD [15].

Figure 2 presents a high-level design framework illustrating how these needs will be addressed through core app features, organized by phase of care. Patient-facing features are shown in dark violet boxes and focus on supporting the lived experience of women navigating cardiovascular conditions, such as personalized education, symptom logging, and wearable-integrated tracking. Clinician-facing features, in the dark gray boxes in the figure, prioritize workflow efficiency, diagnostic support, and data integration. It's also worth noting the interconnectedness of some of these elements. For example, the AI-driven risk alerts and clinical report generation depend on continuous, structured input from patient-generated data, including wearable metrics, symptom tracking, and medication adherence. Similarly, the research and insights feature, designed for clinician relevance, builds on aggregated app data to identify emerging sex-specific patterns and close gaps in clinical evidence.

Figure 2. Design implications across the cardiovascular patient journey in women. AI: artificial intelligence; CVD: cardiovascular disease.



This feature map provides a consolidated foundation for the next phase of development. It will serve as the blueprint for the initial prototype, which will undergo usability testing in an iterative design process, as highlighted in Figure 1 under the methodology section. Through repeated feedback loops with both patients and health care professionals, the prototype will be refined to ensure alignment with real-world needs and system-level constraints, ultimately enhancing adoption, clinical relevance, and long-term impact.

Limitations and Future Research

This study was conducted exclusively in Switzerland, which may limit the generalizability of the findings to other health systems or cultural contexts. Although all patients were recruited via a single university hospital partner, the clinic serves a broad public population and includes individuals with varying levels of health and digital literacy. Nonetheless, this limitation will be addressed in the next phase of the project through iterative prototyping with a more demographically and geographically diverse patient sample. Given that the primary goal is to develop a tool for use in Switzerland, with longer-term plans to expand into other German-speaking countries such as Germany and Austria, we believe the study was appropriately scoped for this stage of development. Broader geographic and cultural representation will be integrated in the next phase of the development, particularly during prototype testing and validation.

The sample size was relatively small, in line with typical qualitative research. While this limits the breadth of perspectives captured, we achieved thematic saturation, suggesting that the findings are sufficiently robust to inform the current design phase. Nonetheless, some stakeholder groups were not represented in this study. In particular, general practitioners and nurses were not included due to recruitment challenges and resource constraints. Their perspectives, especially those of general practitioners, who often serve as the first point of contact during early symptom onset, are crucial for understanding barriers to timely recognition and referral. Future phases of the development will include these stakeholders to ensure that insights from across the care continuum are reflected in the final design.

Conclusions

This study confirmed a persistent disconnect between existing cardiovascular care pathways and the specific needs of women living with CVD. Patients shared experiences of feeling underinformed, underheard, and underserved, particularly during early diagnosis support, due to atypical symptoms and a lack of sex-specific guidance. Clinicians acknowledged these challenges but admitted that current tools and workflows are still largely designed around male-centric risk models. Both groups agreed that digital health has the potential to bridge this gap but also highlighted different expectations; patients seek personalized education and empowerment, while clinicians prioritize clinical use, low burden, and workflow integration.

The participants' input showed strong overlap in desired core features, such as sex-specific symptom and cycle tracking, medication adherence tools, and tailored educational content.

However, divergences appeared around data communication and feedback. Patients favored more interactive, real-time support, while clinicians preferred summarized, actionable insights to avoid alert fatigue and liability concerns. These findings highlight the importance of modular, adaptable design, an app that offers flexibility in use, differentiated interfaces for patients and providers, and clearly defined roles for AI and automation.

These findings also inform key HF considerations for interface design. For clinicians, minimizing cognitive workload was critical; participants favored tools that provide actionable, low-effort insights rather than real-time data streams. Features such as passive data aggregation, structured dashboards, and summary reports were seen as more aligned with clinical workflows and information processing constraints. For patients, usability was tied to personalization, transparency, and flexible communication options. To ensure task-technology fit, the design must allow differentiated engagement based on user role, with modular, customizable features that minimize burden while maximizing relevance.

Success metrics also differed by perspective. Patients emphasized psychological and motivational benefits such as feeling in control, recognized, and informed, whereas clinicians focused on measurable outcomes such as adherence, hospital avoidance, and engagement. However, both agreed that success requires trust, usability, and relevance to daily life. These insights informed a dual-value strategy for app development, aligning patient empowerment with clinical efficiency.

Finally, our expert interviews highlighted that integrating regulatory and reimbursement perspectives early is key to a successful implementation. With features such as AI-supported insights and clinical data sharing, the app will likely qualify as a medical device under Swiss and European laws. Planning for staged implementation, starting with nonregulated educational features and expanding as validation and infrastructure grow, emerges as a practical strategy. However, long-term success and widespread adoption would largely depend on demonstrating real-world value and forming partnerships with insurers and public health programs.

Acknowledgments

The authors thank the Women’s Heart Health Program at the University Hospital Basel for their instrumental role in patient recruitment. Artificial intelligence was not used to create, modify, or interpret the scientific content of the article. It was solely used as a proofreading tool to correct spelling and grammar, enhance clarity, and streamline phrasing. The authors take full responsibility for the integrity and originality of the work.

Funding

This study was sponsored by Innosuisse (Swiss Innovation Agency, grant 75940.1 INNO-ICT). The role of Innosuisse is to promote science-based innovation in the interests of industry and society in Switzerland.

Conflicts of Interest

PV is the founder of Vavken Health Labs, the technology provider developing the application being studied. CJ is an editorial board member of *JMIR Human Factors*, a JMIR Publications journal, at the time of this publication.

Multimedia Appendix 1
Participant information sheet and consent form.
[PDF File (Adobe PDF File), 333 KB - humanfactors_v13i1e82916_app1.pdf]

Multimedia Appendix 2
Sample characteristics of participating patients and clinicians.
[PDF File (Adobe PDF File), 80 KB - humanfactors_v13i1e82916_app2.pdf]

Multimedia Appendix 3
Interview guides.
[PDF File (Adobe PDF File), 123 KB - humanfactors_v13i1e82916_app3.pdf]

Multimedia Appendix 4
Thematic Analysis Process.
[PDF File (Adobe PDF File), 115 KB - humanfactors_v13i1e82916_app4.pdf]

Multimedia Appendix 5
Summary of regulatory and reimbursement considerations.
[PDF File (Adobe PDF File), 95 KB - humanfactors_v13i1e82916_app5.pdf]

References

1. Tromp J, Jindal D, Redfern J, Bhatt A, Séverin T, Banerjee A, et al. World Heart Federation roadmap for digital health in cardiology. Glob Heart 2022;17(1):61 [FREE Full text] [doi: 10.5334/gh.1141] [Medline: 36051317]
2. Leading causes of death. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm [accessed 2025-01-10]
3. Cardiovascular diseases. World Health Organization. URL: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1 [accessed 2025-01-17]
4. Bairey Merz CN, Andersen H, Sprague E, Burns A, Keida M, Walsh MN, et al. Knowledge, attitudes, and beliefs regarding cardiovascular disease in women: the women's heart alliance. J Am Coll Cardiol 2017 Jun 22;70(2):123-132 [FREE Full text] [doi: 10.1016/j.jacc.2017.05.024] [Medline: 28648386]
5. Al Hamid A, Beckett R, Wilson M, Jalal Z, Cheema E, Al-Jumeily Obe D, et al. Gender bias in diagnosis, prevention, and treatment of cardiovascular diseases: a systematic review. Cureus 2024 Feb 15;16(2):e54264 [FREE Full text] [doi: 10.7759/cureus.54264] [Medline: 38500942]
6. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization Institutional Repository for Information Sharing. 2011. URL: https://apps.who.int/iris/handle/10665/44607 [accessed 2022-11-27]
7. Singh S, Kolekar MH, Tanwar S, Wierzchoń ST, Bhatnagar RK, editors. Emerging Technologies for Computing, Communication and Smart Cities. Singapore: Springer Nature Singapore; 2022.
8. Coorey GM, Neubeck L, Mulley J, Redfern J. Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. Eur J Prev Cardiol 2018;25(5):505-521. [doi: 10.1177/2047487317750913] [Medline: 29313363]
9. Castaneda D, Esparza A, Ghamari M, Soltanpur C, Nazeran H. A review on wearable photoplethysmography sensors and their potential future applications in health care. Int J Biosens Bioelectron 2018;4(4):195-202 [FREE Full text] [doi: 10.15406/ijbsbe.2018.04.00125] [Medline: 30906922]
10. Scheid JL, Reed JL, West SL. Commentary: is wearable fitness technology a medically approved device? Yes and no. Int J Environ Res Public Health 2023;20(13) [FREE Full text] [doi: 10.3390/ijerph20136230] [Medline: 37444078]
11. Cowie MR, Lam CSP. Remote monitoring and digital health tools in CVD management. Nat Rev Cardiol 2021:457-458. [doi: 10.1038/s41569-021-00548-x] [Medline: 33824486]
12. Guo Y, Wang H, Zhang H, Liu T, Liang Z, Xia Y, et al. Mobile photoplethysmographic technology to detect atrial fibrillation. J Am Coll Cardiol 2019;74(19):2365-2375 [FREE Full text] [doi: 10.1016/j.jacc.2019.08.019] [Medline: 31487545]
13. Triantafyllidis A, Kondylakis H, Katehakis D, Kouroubali A, Alexiadis A, Segkouli S, et al. Smartwatch interventions in healthcare: a systematic review of the literature. Int J Med Inform 2024;190:105560. [doi: 10.1016/j.ijmedinf.2024.105560] [Medline: 39033723]
14. Jacob C, Lindeque J, Müller R, Klein A, Metcalfe T, Connolly SL, et al. A sociotechnical framework to assess patient-facing eHealth tools: results of a modified Delphi process. NPJ Digit Med 2023;6(1):232 [FREE Full text] [doi: 10.1038/s41746-023-00982-w] [Medline: 38102323]
15. Chauhan GK, Vavken P, Jacob C. Mobile apps and wearable devices for cardiovascular health: narrative review. JMIR Mhealth Uhealth 2025;13:e65782-e65782. [doi: 10.2196/65782]
16. Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ 1995;311(6996):42-45 [FREE Full text] [doi: 10.1136/bmj.311.6996.42] [Medline: 7613329]
17. Mays N, Pope C. Qualitative research in health care. Assessing quality in qualitative research. BMJ 2000;320(7226):50-52 [FREE Full text] [doi: 10.1136/bmj.320.7226.50] [Medline: 10617534]
18. Norman DA, Draper SW. User Centered System Design: New Perspectives on Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates; 1986.
19. The Field Guide to Human-Centered Design. IDEO. URL: https://www.ideo.com/journal/design-kit-the-human-centered-design-toolkit [accessed 2022-06-10]
20. Or CK, Holden RJ, Valdez RS. Human factors engineering and user-centered design for mobile health technology: enhancing effectiveness, efficiency, and satisfaction. In: Human-Automation Interaction: Mobile Computing. Cham, Switzerland: Springer; 2023:97-118.
21. 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;9(12):e28102 [FREE Full text] [doi: 10.2196/28102] [Medline: 34874893]
22. Jacob C, Bourke S, Heuss S. From testers to cocreators—the value of and approaches to successful patient engagement in the development of ehealth solutions: qualitative expert interview study. JMIR Hum Factors 2022;9(4):e41481. [doi: 10.2196/41481]
23. Nabid MR, Sedghi R, Behbahani M, Arvan B, Heravi MM, Oskooie HA. Application of poly 1,8-diaminonaphthalene/multiwalled carbon nanotubes-COOH hybrid material as an efficient sorbent for trace determination of cadmium and lead ions in water samples. J Mol Recognit 2014;27(7):421-428. [doi: 10.1002/jmr.2361] [Medline: 24895274]

24. Creswell J, Poth C. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: SAGE Publications, Inc; 2017.
25. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Laborante R, Borovac JA, Galli M, Rodolico D, Ciliberti G, Restivo A, et al. Gender-differences in antithrombotic therapy across the spectrum of ischemic heart disease: time to tackle the Yentl syndrome? Front Cardiovasc Med 2022;9 [FREE Full text] [doi: 10.3389/fcvm.2022.1009475] [Medline: 36386309]
28. de Marvao A, Alexander D, Bucciarelli-Ducci C, Price S. Heart disease in women: a narrative review. Anaesthesia 2021;76 Suppl 4:118-130 [FREE Full text] [doi: 10.1111/anae.15376] [Medline: 33682102]
29. Keteepe-Arachi T, Sharma S. Cardiovascular disease in women: understanding symptoms and risk factors. Eur Cardiol 2017;12(1):10-13 [FREE Full text] [doi: 10.15420/ecr.2016:32:1] [Medline: 30416543]
30. Panahiazar M, Bishara AM, Chern Y, Alizadehsani R, Islam SMS, Hadley D, et al. Gender-based time discrepancy in diagnosis of coronary artery disease based on data analytics of electronic medical records. Front Cardiovasc Med 2022;9 [FREE Full text] [doi: 10.3389/fcvm.2022.969325] [Medline: 36505372]
31. Park LG, Kariuki JK, Hendriks JM, Gallagher R, Burke LE. A review and promise of digital health technologies to promote global cardiovascular health: a call to action for nursing. J Cardiovasc Nurs 2025. [doi: 10.1097/JCN.0000000000001205] [Medline: 40241412]
32. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding clinicians' adoption of mobile health tools: a qualitative review of the most used frameworks. JMIR Mhealth Uhealth 2020;8(7):e18072 [FREE Full text] [doi: 10.2196/18072] [Medline: 32442132]
33. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. JMIR Mhealth Uhealth 2020;8(2):e15935 [FREE Full text] [doi: 10.2196/15935] [Medline: 32130167]
34. Jacob C, Sezgin E, Sanchez-Vazquez A, Ivory C. Sociotechnical factors affecting patients' adoption of mobile health tools: systematic literature review and narrative synthesis. JMIR Mhealth Uhealth 2022;10(5):e36284 [FREE Full text] [doi: 10.2196/36284] [Medline: 35318189]
35. Jacob C, Brasier N, Laurenzi E, Heuss S, Mougiakakou S, Cöltekin A, et al. AI for IMPACTS framework for evaluating the long-term real-world impacts of AI-powered clinician tools: systematic review and narrative synthesis. J Med Internet Res 2025;27:e67485 [FREE Full text] [doi: 10.2196/67485] [Medline: 39909417]
36. Piani F, Baffoni L, Strocchi E, Borghi C. Gender-specific medicine in the european society of cardiology guidelines from 2018 to 2023: where are we going? J Clin Med 2024;13(14) [FREE Full text] [doi: 10.3390/jcm13144026] [Medline: 39064064]
37. Gill SK, Barsky A, Guan X, Bunting KV, Karwath A, Tica O, et al. Consumer wearable devices for evaluation of heart rate control using digoxin versus beta-blockers: the RATE-AF randomized trial. Nat Med 2024;30(7):2030-2036. [doi: 10.1038/s41591-024-03094-4] [Medline: 39009776]
38. Fomo M, Borga LG, Abel T, Santangelo PS, Riggare S, Klucken J, et al. Empowering capabilities of people with chronic conditions supported by digital health technologies: scoping review. J Med Internet Res 2025;27:e68458 [FREE Full text] [doi: 10.2196/68458] [Medline: 40577044]
39. Avşar TS, Elvidge J, Hawksworth C, Kenny J, Németh B, Callenbach M, et al. Linking reimbursement to patient benefits for advanced therapy medicinal products and other high-cost innovations: policy recommendations for outcomes-based agreements in Europe. Value Health 2024;27(11):1497-1506 [FREE Full text] [doi: 10.1016/j.jval.2024.07.007] [Medline: 39094693]
40. van Kessel R, Srivastava D, Kyriopoulos I, Monti G, Novillo-Ortiz D, Milman R, et al. Digital health reimbursement strategies of 8 European countries and Israel: scoping review and policy mapping. JMIR Mhealth Uhealth 2023;11:e49003 [FREE Full text] [doi: 10.2196/49003] [Medline: 37773610]
41. 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;10:e45143 [FREE Full text] [doi: 10.2196/45143] [Medline: 36843321]
42. Fraser AG, Biasin E, Bijnens B, Bruining N, Caiani EG, Cobbaert K, et al. Artificial intelligence in medical device software and high-risk medical devices - a review of definitions, expert recommendations and regulatory initiatives. Expert Rev Med Devices 2023;20(6):467-491 [FREE Full text] [doi: 10.1080/17434440.2023.2184685] [Medline: 37157833]
43. Factsheet on the reimbursement of digital health applications under the mandatory health insurance (OKP) [Report in German]. Federal Office of Public Health (FOPH). Bern, Switzerland: Bundesamt für Gesundheit (BAG); 2024. URL: https://backend.bag.admin.ch/fileservice/sdweb-docs-prod-bagadminch-files/files/2025/03/18/00d21ad2-1f9d-4c04-803b-622ea70a127f.pdf [accessed 2025-08-06]
44. Arcà E, Heldt D, Smith M. Comparison of health technology assessments for digital therapeutics in Germany, the United Kingdom and France. Digit Health 2025;11 [FREE Full text] [doi: 10.1177/20552076241308704] [Medline: 39845519]

45. Gehder S, Goeldner M. Unpacking performance factors of innovation systems and studying Germany’s attempt to foster the role of the patient through a market access pathway for digital health applications (DiGAs): exploratory mixed methods study. J Med Internet Res 2025;27:e66356. [doi: 10.2196/66356]
46. Wang LY, Chiang GSH, Wee CF, Chan SWK, Lau JXX, Taeihagh A. Preventing ischemic heart disease in women: a systematic review of global directives and policies. npj Womens Health 2024;2(1). [doi: 10.1038/s44294-024-00040-0]

Abbreviations

AF: atrial fibrillation
AI: artificial intelligence
BP: blood pressure
CE: Conformité Européenne
CVD: cardiovascular disease
EHR: electronic health record
EU: European Union
HCD: human-centered design
HF: human factors
HR: heart rate
MDR: Medical Device Regulation
mHealth: mobile health

Edited by J Sarvestan; submitted 25.Aug.2025; peer-reviewed by SCL Au, V P?rez Jover; comments to author 24.Nov.2025; revised version received 12.Dec.2025; accepted 15.Dec.2025; published 13.Jan.2026.

Please cite as:
Jacob C, Puthanveettil SR, Vavken P, Kaplan E, Zuern CS
Bridging Gaps in Women’s Heart Health: User-Centered Needs Assessment Informed by Patient and Clinician Interviews
JMIR Hum Factors 2026;13:e82916
URL: https://humanfactors.jmir.org/2026/1/e82916
doi:10.2196/82916
PMID:

Original Paper

Exploring Video Consultations Across the Public and Private Sectors in Norway: Semistructured Interview Study

Mari Skoge[1,2], ClinPsy; Sofie Ragnhild Aminoff[1], ClinPsy, PhD; Henrik Myhre Ihler[1], MD, PhD; Kari Jorunn Kværner[1,3], MD, PhD; Linn Nathalie Støme[1], MS, PhD; Kristin Lie Romm[1,2], MD, PhD

[1]Early Intervention in Psychosis Advisory Unit for South East Norway, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

[2]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

[3]Department of Strategy and Entrepreneurship, BI Norwegian Business School, Oslo, Norway

Corresponding Author:
Mari Skoge, ClinPsy
Early Intervention in Psychosis Advisory Unit for South East Norway
Division of Mental Health and Addiction
Oslo University Hospital
Sognsvannsveien 21, Bygg 12, 2. etg
Oslo, 0372
Norway
Phone: 47 93624119
Email: maskoge@uio.no

Abstract

Background: Delivering therapy through video consultations can increase the reach and impact of mental health care services. However, adoption varies, and there is a lack of professional consensus about the usefulness of video consultations in therapy settings.

Objective: This study aimed to explore mental health professionals' experiences with and attitudes toward video consultations across different clinical environments in the private and public health care sectors in Norway to inform the design of future digitalized services.

Methods: In this qualitative study, we recruited leaders and clinicians from public hospitals and private clinics. We conducted semistructured interviews that mapped individual experiences and attitudes concerning video consultations, as well as contextual aspects concerning the participants' professional environments. We used reflexive thematic analysis with an inductive, essentialist, and experiential orientation to analyze the data.

Results: A total of 24 mental health professionals (16 from public hospitals and 8 from private clinics) participated. Variations in their attitudes did not follow patterns reflecting the type of service or sector they worked in. Rather, attitudes seemed related to higher-level assumptions rooted in professional culture, societal values, and previous experiences. We generated six themes capturing and structuring the professional perspectives: (1) meta-perspectives on the digitalization of therapeutic rooms, (2) the "how" of service integration, (3) challenging therapist culture, (4) negotiating the limits of the digital therapy room, (5) creating clinical value from the digital format, and (6) adapting techniques and technology in digital therapy sessions.

Conclusions: To strengthen the adoption and impact of video consultations, we should direct attention toward higher-level societal and cultural aspects that shape attitudes and practices. We suggest incorporating digitalized therapy in education, facilitating personal experiences with video consultations, increasing the sharing of knowledge between clinical environments, and sparking innovation of both service models and technology.

(JMIR Hum Factors 2026;13:e80812) doi:10.2196/80812

KEYWORDS

clinician; hybrid psychotherapy; private sector; public sector; telemental health; telepsychiatry; therapist; video consultations; videoconferencing

Introduction

Background

Implementation of video consultations (VC) as an alternative for mental health care delivery addresses several practical barriers to patients' access to treatment [1-4]. The clinical outcomes of therapy delivered through VC are comparable to those of traditional in-person therapy for a wide range of mental health conditions, including depression, suicidal ideation, generalized anxiety disorder, and posttraumatic stress disorder [5-9]. Thus, studies investigating adaptations of psychotherapy practices to the VC format have emerged [10-14]. The literature on VC in services for severe mental illness and other less common mental health conditions is scarcer and inconclusive, but remains optimistic and encourages further research [15-19].

Hybrid therapy—the combination of VC and in-person therapy sessions—allows patients and clinicians to experience the benefits of both formats [20-24]. Previous research on the clinical use of VC has largely focused on comparing the service delivery alternatives in isolation. However, the flexible integration of both VC and in-person treatment modalities is increasingly treated as a promising approach for future mental health care services, potentially providing added value compared to using a single format [25-27]. Studies supporting hybrid models for delivering different types of clinical interventions have appeared in the literature [28-30], but there is a need for more research to develop ideal strategies for integrating VC as a component in routine hybrid services.

The adoption of VC varies significantly across mental health care [31,32], and the attitudes of mental health professionals contribute to this variation [33-36]. Some studies have shown that mental health professionals have largely positive attitudes toward VC, perceiving the format as acceptable and feasible for several clinical purposes [2,33,34,37-41]. Other studies have found greater diversity in perceptions of VC [42,43], and some have identified fundamental drawbacks reported by mental health professionals, such as feeling less present and experiencing a reduction in the quality of the therapeutic relationship [44-46]. Across studies, many report a lack of VC competency or an unmet need for support [14,35,36,47,48]. Furthermore, more experience with VC is associated with attitudes that are more positive [49-51]. Overall, mental health professionals show a preference for the traditional format [24,33,39], which is more evident than that of patients [37,48].

Organizational context is another factor that has been suggested as significant in influencing mental health professionals' attitudes toward and intention to use VC [35]. The public and private health sectors represent distinct organizational conditions. In Norway, the number of private actors offering on-demand mental health care services has increased during the last decade, typically including digital treatment alternatives and opportunities for quicker access to services [52]. In parallel, the proportion of the population with private health insurance (as a supplement to the basic public coverage) has grown significantly [53], which influences the demand in the market for private health services. This development might indicate a more innovative climate, higher responsiveness to user demands, and quick adaptations to digitalization trends among private practitioners. However, we know little about how the private mental health sector approaches VC [54-56]. Health care systems are generally complex and characterized by organizational silos that challenge knowledge sharing across clinical environments [57-59]. Such gaps are highly noticeable within the public health care sector but are even more evident between the public and private sectors [60,61]. Another important reason for the lack of knowledge about private mental health actors is that the private sector in Norway typically has fewer incentives to spend resources on research activities, while the public sector finances and conducts most of the studies on mental health care services.

Study Objectives

In summary, the topic of VC in mental health care lacks professional consensus. In this study, we aim to explore the experiences, attitudes, and contextual aspects related to mental health professionals' use of VC. The study will include professionals working in various services across public and private mental health care to capture a broad range of perspectives. Our ambition is to develop knowledge about factors that contribute to differences in VC adoption and attitudes to inform the design of accessible future health care services that reach more people.

Methods

Study Design

This study is a qualitative investigation of mental health professionals' experiences with and attitudes toward VC, as well as the professional contexts of the participants. The study uses a combination of convenience and snowball sampling strategies [62], digital semistructured interviews, and reflexive thematic analysis [63].

Study Setting

The study was conducted in Norway. The Norwegian national public health system provides universal health coverage. The private health expenditure of the population is generally low (14% in 2021) [64]. Mental health care is provided through primary and specialized services, the latter requiring the presence of moderate or severe psychiatric conditions and a referral from the primary services. Specialized psychiatric services comprise both district psychiatric centers (local outpatient clinics) and hospitals. The services are semidecentralized and organized based on catchment areas, aiming to provide equal access to care across all social strata and geography [65].

The 2 public sector sites of our study are St Olav's Hospital (Trondheim University Hospital) and Oslo University Hospital. The catchment area of St Olav's Hospital consists of both urban and rural areas in Central Norway and is characterized by greater geographical distances between inhabitants and services than the catchment area of Oslo University Hospital, which serves urban areas only. The 4 private sector sites of the study are all located in major cities of Norway. These private clinics provide services to people who are self-referred. Their clients can pay for the services themselves or through private health insurance.

Three of the private clinics in our study allow clients to book consultations directly through their websites, where clients can access information about available therapists and time slots. The fourth clinic requires clients to call, email, or submit a contact form. Two of the private clinics specialize in certain therapy orientations and patient populations, while the others offer more general mental health services. To protect the anonymity of the participants, we have chosen not to disclose the names of these clinics, as they are small, and the mental health professionals could be identified.

Recruitment

The sample of participants in this study was obtained through convenience sampling, which included a snowball strategy [62]. We contacted the leaders of psychiatric clinics at St Olav's Hospital, Oslo University Hospital, and several private clinics and invited them to take part as recruitment sites in our study (see Figures S1-S3 in Multimedia Appendix 1 for more details). The 2 public hospitals and 4 of the private clinics agreed to participate and approved the recruitment and data collection procedures. Two private clinics did not respond to our invitations. Both management and study participants at the clinical sites facilitated further recruitment of participants by informing colleagues about the project and sharing the project group's contact information.

The inclusion criteria were as follows: (1) working as a clinician or working as a leader of other clinicians and (2) employed at one of the public hospitals or affiliated with one of the private clinics. All heads of the units had approved their participation and accepted that the participants spent an hour of their day on the interview. We did not aim to achieve a representative sample to generate representative findings; rather, we aimed to develop empirical thematic insights that were transferable to both research and clinical contexts [66]. We sought to access a broad range of viewpoints and contexts that would allow an exploration of whether expected characteristics (eg, sector affiliation) and other nonpredefined aspects influence approaches to VC. Importantly, we also included participants working in services targeting severe mental illnesses. A lesson learned from our previous work is that researchers and services should consider patients living with severe diagnoses as candidates for digitalized services, although they are often excluded from the target groups of such interventions [67].

Data Collection

Data collection was carried out between May 2024 and December 2024. We conducted digital semistructured interviews addressing individual experiences and attitudes toward VC, as well as the participants' workplace and professional environment. The interviews were based on an interview guide developed by the authors for the purpose of this study (Multimedia Appendix 2). We used the secure web-based VC platform Join, delivered by the state-owned company Norsk helsenett. The duration of the interviews ranged from 45 to 60 minutes. We recorded each interview with an app developed by Services for Sensitive Data (TSD), owned by the University of Oslo. The interviews were automatically transferred to the secure TSD platform and transcribed by the University of Oslo's OpenAI Whisper technology.

Data Analysis

We analyzed the interview transcripts from a critical realist position, applying a method largely based on Braun and Clarke's [63] reflexive thematic analysis. We approached the data with an inductive, explorative, experiential, and contextualist approach [68]. The analysis was data-driven. It placed a direct focus on the participants' utterances rather than on language and latent meaning. A contextualist orientation was a natural choice to capture the features of professional cultures and clinical environments. Choosing the reflexive school of thematic analysis, we aimed to transparently communicate the impact of researcher participation, as well as analytic decisions and processes. We selected this method with the ambition of developing meaningful themes that can carry transferable insights across the mental health field [66]. Moreover, the analysis was conducted from a pragmatic perspective that considered current and future clinical practices. We followed the 6 phases of thematic analysis outlined by Braun and Clarke [69], with some minor adaptations (Multimedia Appendix 3).

User Representation

Two user representatives with lived experience, recruited through the Bipolar Association Norway, contributed to this study. The representatives provided feedback on the study in the early design phase and during the writing process. The purpose of this involvement was to ensure that the research aims, the methods applied to achieve these, and the angle used to present and discuss the findings are relevant, interesting, and valuable to people living with mental disorders, including severe conditions such as bipolar and psychotic disorders.

Ethical Considerations

The project obtained approval from the data protection officer at Oslo University Hospital (24/04658). All participants provided written consent to take part in the study. The consent forms included information about the purpose of data collection and how the data would be stored and processed. The forms emphasized voluntary participation, including the option to leave the study at any time without risking any consequences. Signed consents and interview transcripts were stored in the secure TSD database. To protect the privacy of participants, identifiable features of participants were removed by MS within the secure TSD environment before the data were exported for analysis.

The economic compensation offered only to the participants working in private clinics entails ethical consideration. Importantly, private practitioners' financial models conflict with involvement in research, which contributes to the maintenance of a knowledge gap concerning private health care services. Moreover, the project group holds the view that the compensation was not high enough to create any form of pressure for the private practitioners to participate in the study, only enough to compensate for lost income during the interview. Therefore, we evaluate the economic compensation to the participants working in private clinics as both ethically acceptable and reasonable.

Results

Overview

A total of 24 mental health professionals participated in our study (Figure 1 and Table 1), representing different demographic characteristics, geographical catchment areas, professional backgrounds, types of services, and target patient populations, as well as affiliation to either the public or private health sector. From the earliest phases of the analysis, we observed great variation in the participants' attitudes toward VC and in the amount of personal experience with VC as a therapy format. Figure 2 is a result of early familiarization with the data material (Multimedia Appendix 3). The figure provides a simplified overview of the 3 subgroups in regard to levels of VC experience and attitudes. The figure is based on subjective and exploratory screening of the main tendencies in the participant accounts. The most extreme values on each of the axes represent the most extreme participant views identified in this study sample, and the position of each data point is relative to the others. Taken together, the distribution of data points in the figure suggests that there is a tendency that VC experiences and VC attitudes are related and that the relationship between the two concepts is stronger at more extreme levels. Moreover, the figure indicates differences in the amount of experience with VC between the 3 subgroups. Nevertheless, a continuum of attitudes is represented in each group. Early screening of attitudes in relation to participants' characteristics indicated that variations in attitudes did not systematically follow affiliation with the public or private sector or the type of service the mental health professionals worked with.

Figure 1. Chart of recruitment sites and study participants.

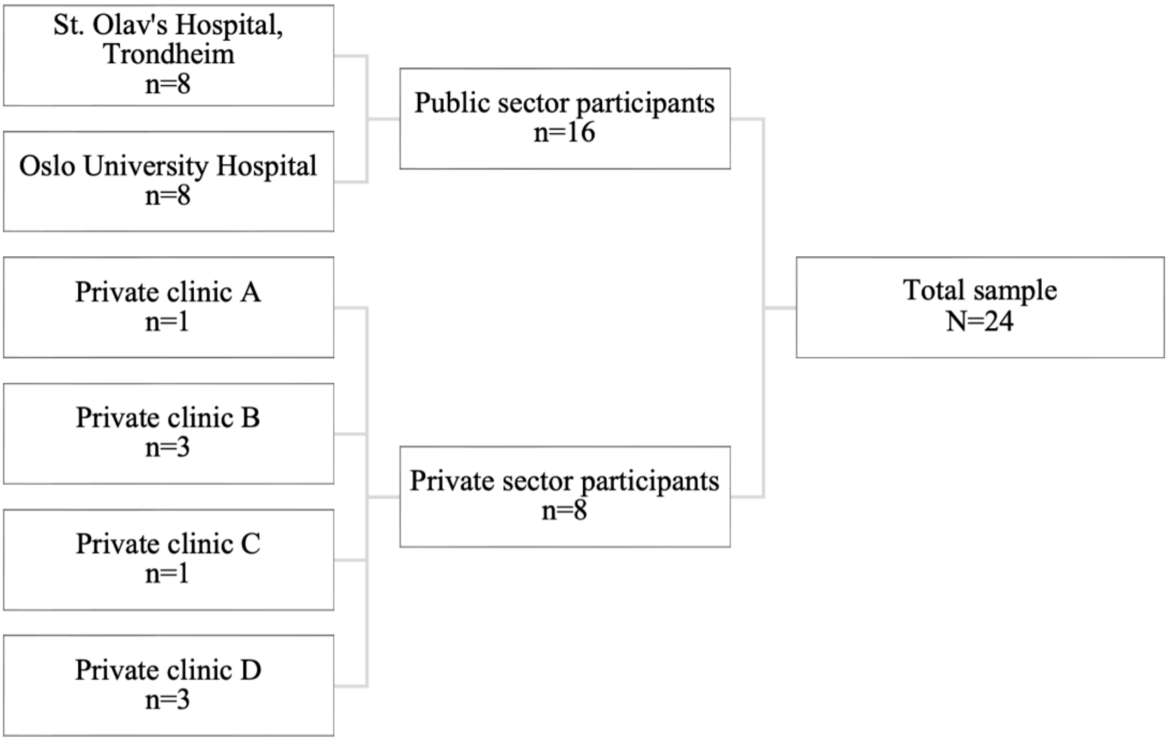
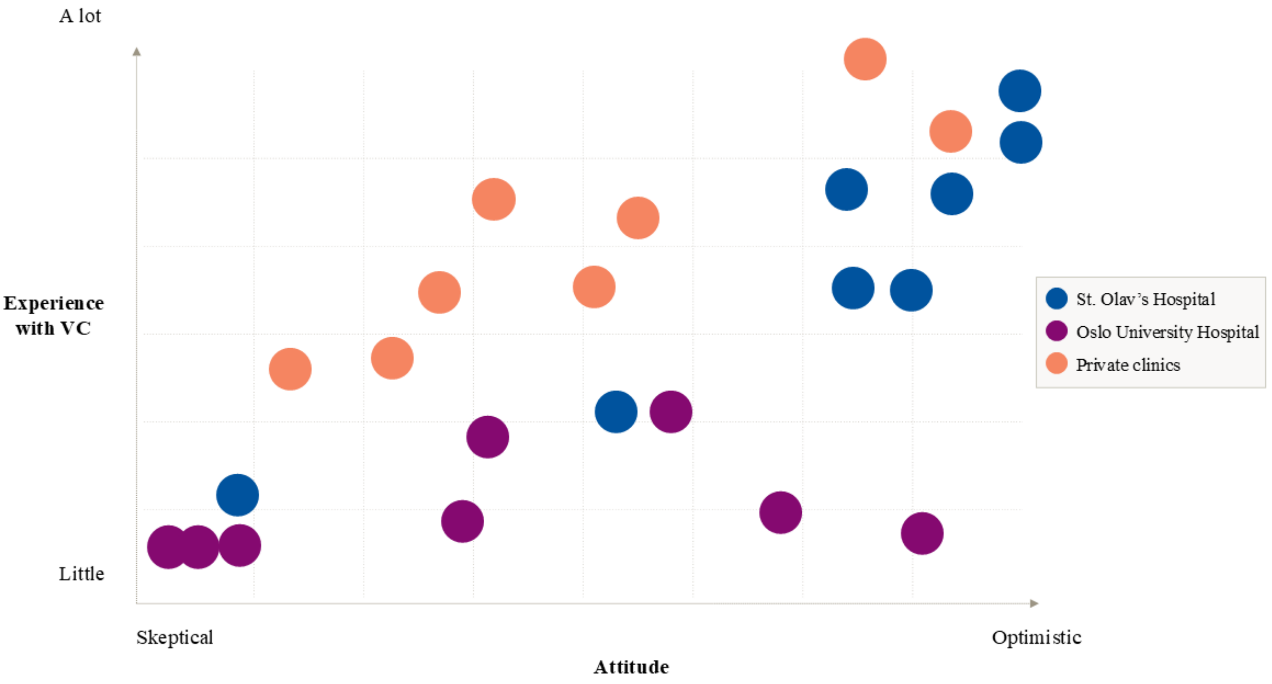


Table 1. Participant characteristics (N=24).

Characteristics	Values
Participant age (years), mean (SD; range)	45.5 (9.9; 31-65)
Sex assigned at birth, n (%)	
Female	17 (71)
Male	7 (29)
Role, n (%)	
Clinician	16 (67)
Leader (few clinical tasks in everyday workflow)	3 (12)
Combined (significant amount of clinical tasks in everyday workflow)	5 (21)
Professional background, n (%)	
Specialist in clinical psychology	9 (38)
Licensed clinical psychologist	5 (21)
Psychiatrist	4 (17)
Resident doctors in psychiatry	3 (12)
Psychiatric nurse	2 (8)
Nurses	1 (4)
Type of mental health service, n (%)	
Open booking/no referral	8 (33)
Psychotic disorders	5 (21)
General psychiatric services	4 (17)
Short-term general psychiatric services	3 (12)
Bipolar disorder	3 (12)
Currently not working in a clinical service	1 (4)

Figure 2. Early qualitative mapping of the experiences and attitudes of participants. VC: video consultations.



Through the following phases of the reflexive thematic analysis, we constructed a set of 6 themes, with each theme belonging to a distinct level of abstraction (Textbox 1).

Textbox 1. Themes from the reflexive thematic analysis listed according to level.

Meta level

- Theme 1: meta-perspectives on the digitalization of therapeutic rooms

Service level

- Theme 2: the "how" of service integration

Therapist culture level

- Theme 3: challenging therapist culture

Clinical level

- Theme 4: negotiating the limits of the digital therapy room
- Theme 5: creating clinical value from the video format

Practice level

- Theme 6: adapting techniques and technology in digital therapy

Theme 1: Meta-Perspectives on the Digitalization of Therapeutic Rooms

The first theme connects the practice of VC to the overarching digitalization trends in society and health care services. The participants demonstrated opposing underlying values that naturally influenced their standpoints at lower levels (ie, the 5 other themes). Interestingly, these meta-perspectives did not seem to follow other key characteristics of the participants, such as age, gender, private or public workplace, clinic, or severity of the condition of their patient population.

Some viewed the increased attention paid to VC in mental health care as a symptom of the fast-paced digitalization of our communities, which reduces arenas for healthy ways of experiencing and relating to others. Using this rationale, VC practices contribute to an unwanted development toward more individualism and isolation. Thus, there is a need to safeguard some of these analogue spaces, especially traditional therapy rooms, which might be particularly important.

> *Personally, and on behalf of others, I find it very exhausting to interact so much digitally. It has an overhead cost, whose extent I think we don't fully understand yet.* [B4, resident doctor in psychiatry, public hospital]
>
> *Digital (encounters) enhance the extremely rational, intellectualizing way we relate to ourselves and others, which I believe results in more psychopathology.* [D1, specialized psychologist, private clinic]

The opposite viewpoint places digitalization in relation to new dynamics in the distribution of power and agency between patients and services. Since offering digital treatment can increase access to care and help balance the relational asymmetry of health care services, this viewpoint treats the expansion of therapy alternatives as a moral obligation.

> *They experience a higher degree of autonomy and control when they meet me at a screen's distance.* [A1, psychiatrist, public hospital]
>
> *We deprive some patients of the opportunity to show up—patients who would otherwise struggle because they must enter our arena. There is already an imbalance—an asymmetrical relationship. Therefore, I think we owe it to them.* [A2, specialized psychologist, public hospital]

Theme 2: The "How" of Service Integration

The second theme captures the diversity in VC integration strategies and emphasizes the relevance of exploring creative possibilities of VC use further. The data material contained descriptions of various service models involving VC in different ways, serving different purposes. See Textbox 2 for an overview of the service models identified in the study. The list demonstrates that using VC does not mean one thing. Rather, VC can be adapted to the specific needs of different clinical contexts. In practice, the participants often applied a combination of elements from the listed service models. We identified the most evident differences between the sectors in these dimensions. Some models were more available or exclusively implemented in the private sector, such as the digital hub of specialists. The hub allows personalization on several levels.

Textbox 2. Service models that integrate videoconferencing identified in the data material.

Only digital

- All consultations in a course of treatment are digital. One private therapist worked exclusively with video consultations (VC) from home. One public clinician worked with a fully digital intervention for relatives.

Digital first session

- All new patients attend their first session through VC. Later, they decide on the format. Used in a public service for bipolar disorder that served many patients referred directly from emergency wards.

Hybrid strategy (relational booster rationale)

- Hybrid courses of treatment where traditional consultations comprise a minority of meetings and are used to strengthen the therapeutic relationship. Patients typically live far away.

Hybrid strategy (content rationale)

- Hybrid courses of treatment where the format depends on the session's content. For instance, parts of psychiatric assessments where only verbal reports are needed can indicate a VC session.

Hybrid strategy (practical rationale)

- Hybrid courses where there is a preplanned strategy of using both formats to reduce travel or absence from daily activities. Target patients may have a network and activities they attend regularly.

Hybrid strategy (symptom-based rationale)

- Hybrid courses where anxiety-related symptoms are barriers to receiving services, and therapy starts off as VC, while the proportion of physical meetings gradually increases based on symptom reduction.

Hybrid strategy (staff flexibility rationale)

- Clinicians work from home on days when only digital appointments are scheduled to increase variety, flexibility, and satisfaction with work.

Ad hoc hybrid courses (short-notice changes)

- Traditional sessions may change into digital ones at short notice due to patients experiencing unanticipated events or problems with executive functions. VC sessions secure continuity.

Open booking (only private clinics)

- A client-centered service that allows anyone to book a VC session at a suitable time with the therapist they prefer, typically directly through a website. Some clinics also offer different durations of sessions.

Digital hub (only private clinics)

- Clinics provide information about therapists' competencies and preferred therapy orientations, offering a menu of therapeutic alternatives from which clients can choose the best match for their needs.

> *Here, you can choose someone who matches where you are in the therapy process, what your goals are, and how you like to work. I think this leads to faster improvement compared to going to someone at random.* [D1, specialized psychologist, private clinic]

Furthermore, working independently with more flexible opening hours in a private clinic allows more staff flexibility for practitioners:

> *It is nice to have some days at home, deciding for myself how I want to structure the day. I might have a consultation at nine, the next consultation at one. Then one at six, and the last one at eight. Depending on what the client needs. And meanwhile, I can just live my life.* [E2, psychologist, private clinic]

Hybrid, preplanned strategies for delivering therapy for practical and symptom-related reasons were applied across the sample. However, a few therapists also talked about using traditional sessions to boost the relational quality of mainly digital courses of treatment, where the effect of the physical conversation would be maintained for a while:

> *... then we can keep leveraging from that (physical) booster session in the other (digital) conversations.* [E3, specialized psychologist, private clinic]

An insight from this dimension was that the clinicians' presentation of how a VC service could work out is important, as it can shape patients' expectations. One of the participants worked with a completely digital support service for the relatives of patients with severe mental illness. The clinician described how the relatives accepted the digital service they were offered and that there were no requests for in-person meetings:

> *The thing about the (digital service to relatives) conversations is that they are video-only. Right? Even*

> *from the first session. So, you go into it with the expectation that we will only meet on video.* [B5, psychiatric nurse, public hospital]

Theme 3: Challenging Therapist Culture

The third theme is based on the participants' descriptions of VC as a deviation from tradition and something that challenges assumptions established in professional culture. Some participants described video practice as a threat to therapists' integrity, associating it with something less therapeutic than traditional therapy. Others said that VC could trigger a feeling of helplessness and less confidence in their role as therapists, a threat to their professional identity.

> *It's kind of taboo in a way ... I think some view it as somewhat poorer. That you don't really connect with the patient properly, when you don't meet and look this person properly in the eyes.* [B7, nurse, public hospital]

> *There is sometimes a need to initiate acute interventions, and then you don't necessarily get in the right position to act ... you feel more helpless on video.* [B2, resident doctor in psychiatry, public hospital]

Nevertheless, several participants challenged their preconceptions through experience and developed new, accepting approaches to VC. They reflected on the norms and myths of psychotherapy and identified culture-based assumptions about digital therapies:

> *If people were used to sitting at home and talking, and that was the standard, I am pretty sure going to an office would have felt strange.* [A6, specialized psychologist, public hospital]

> *It has become imperative that the patient comes to the outpatient clinic and stays there for 45 minutes. I do not have any evidence to state that it is very effective or useful for either the patient or us. ... It is deeply ingrained in healthcare culture. Things like video consultations are not part of education programs. ... This influences what one thinks as a clinician.* [A7, specialized psychologist, public hospital]

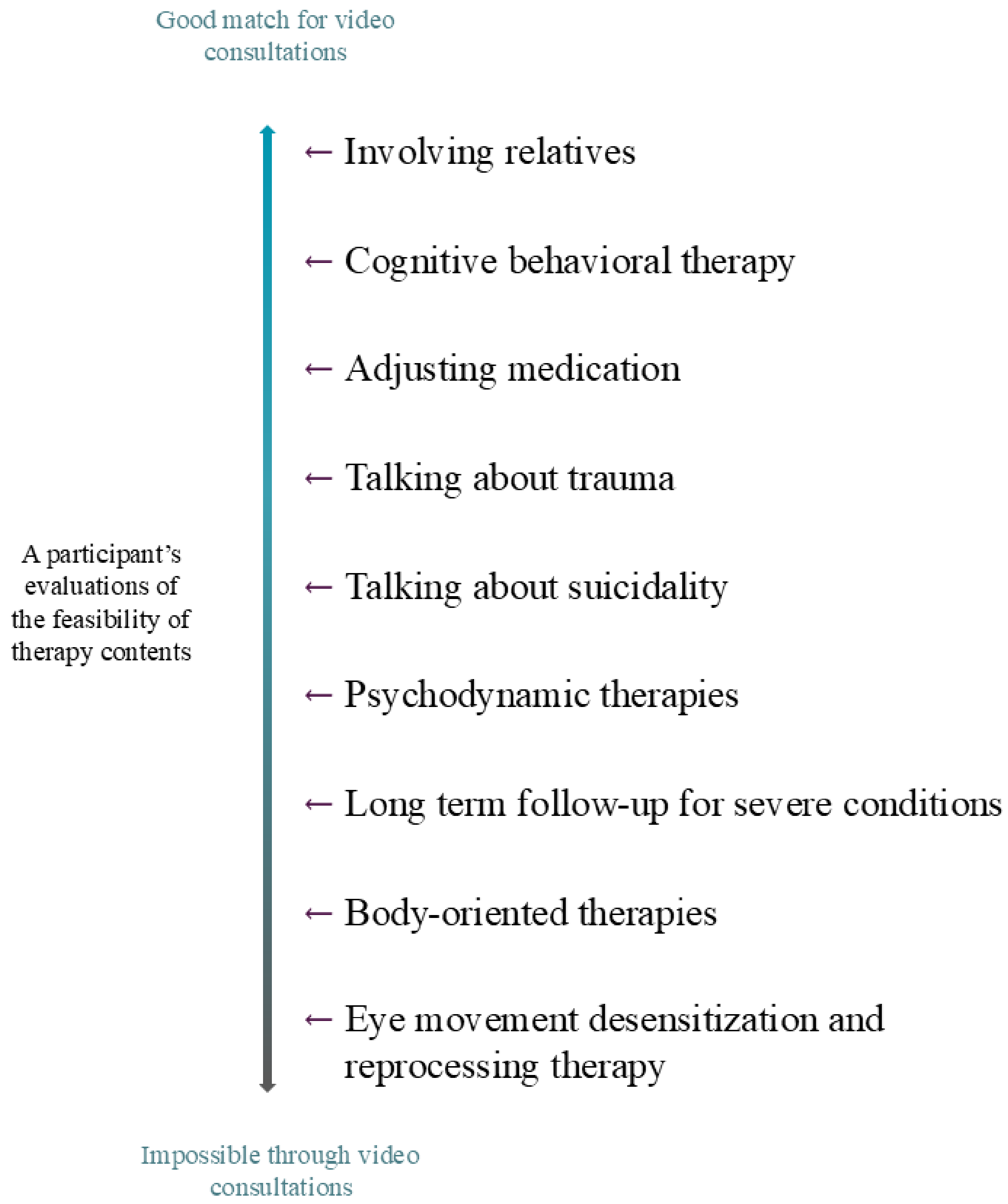
As observed early in the familiarization process of the dataset, there seemed to be a trend that the most experienced VC therapists held more positive views, and vice versa. However, we also observed that private practitioners with more experience than employees from public hospitals conveyed diverse opinions. Interestingly, in private clinics, patients decide on the therapy format themselves when they book a consultation. Thus, some private practitioners may practice an even amount of VC therapy, even if they subscribe to a therapist culture that favors the physical therapy space.

Theme 4: Negotiating the Limits of the Digital Therapy Room

Theme 4 demonstrates the contrasts in professional opinions and rationales regarding the clinical purposes for which VC is suitable. All participants acknowledged that the digital therapy room was less appropriate when patients lacked social contact and activities in their daily lives or needed exposure to work with avoidance. Apart from this, their reflections on the feasibility of different therapy contents varied.

Figure 3 illustrates the continuum of the perceived feasibility of different therapy contents when delivered through VC, based on a participant who worked within the cognitive therapy orientation. The case presented in the figure demonstrates how participants typically had clear logic and presented a continuum of feasibility, defining the limits of VC. From the viewpoint in this case, eye movement desensitization and reprocessing therapy (EMDR) is the least feasible intervention using VC.

Figure 3. Illustration of a typical continuum of feasibility concerning different therapy contents when delivered through videoconferencing, based on 1 participant.



However, others expressed contrasting opinions. These perspectives on feasibility are demonstrated in the following extracts:

> *There are certain things that are not easy to do on video—that cannot be done on video. For example, trauma therapy with EMDR. It is completely impossible.* [A3, psychiatrist, public hospital]
>
> *I might think that EMDR works better digitally than other forms of therapy, because there are very few words.* [D1, specialized psychologist working with EMDR in traditional sessions, private clinic]

Similarly, while most participants talked about patients with psychotic disorders as the patient population least probable of experiencing value from VC, one of the clinicians working with this patient group remarked,

> *I think the patient can often surprise the clinician in that area, and they can often be better (than what the clinicians expect).* [B7, nurse, public hospital]

Finally, some mental health professionals demonstrated the importance of understanding the difference between the feasibility of carrying out interventions and the actual therapeutic value that is realized. In the context of services aimed at patients struggling with severe diagnoses and complex life challenges, a VC may not achieve the same outcomes as in-person contact. A participant stated that services and clinicians are at risk of misleading themselves in the following way:

> *"Now that we have digital tools, we actually get in touch with the patient. The patient attends the appointments. We are now in a position where we can talk to this person." However, the patient just sits there in their municipal apartment, isolated,*

> *without needing to go out. "But at least we get in touch with him!" I imagine that we might end up just believing that we have done something good.* [B4, resident doctor in psychiatry, public hospital]

Theme 5: Creating Clinical Value From the Video Format

The fifth theme revolves around the process of realizing therapeutic value from VC. The data material contained examples demonstrating that the video format, in some cases, comes with unique additional clinical benefits:

> *I have many clients whom I have never met in person. However, it was particularly interesting that with those I had met in person, I felt that more happened when we could talk on video, and I experienced more therapeutic breakthroughs with them after we started with video consultations. I found that a bit interesting.* [E2, psychologist, private clinic]

Several clinicians explained that the practical benefits of the video format mediated some of the extra therapeutic value. When patients do not need to spend much time travelling or spend energy overcoming practical obstacles and stress related to a physical meeting, more resources may be available for the purpose of the therapy session. Moreover, remote sessions often come with proximity to family members, which can result in more effective therapeutic processes:

> *They might have to sit on a bus for two hours each way, and then the entire consultation would end up talking about how terribly frustrating it had been. When you have a video consultation, you can really concentrate on what the underlying problem is.* [A3, psychiatrist, public hospital]

> *I had the opportunity to do things more quickly. I found out that it might be a good idea to talk to (the patient's) relative just to calm the situation and asked if that relative was present at home. The patient said yes, and then I got 15 minutes with the relative right away. And then that situation was solved. It wasn't an issue anymore.* [A6, specialized psychologist, public hospital]

Another application of VC is as an aid to regulate patients' level of activation. Facilitating a safe, less confronting therapy environment reduces the level of activation, providing patients with more opportunities to take part in and experience the effects of emotionally challenging therapeutic interventions:

> *If there is a really difficult conversation with a client where we are going to talk about our alliance, I might take the client out of the therapy room to help downregulate feelings, not having to maintain eye contact. We can use video to do the same. Sometimes, when we want feelings to be more intense, I prefer to be in the office.* [E1, psychologist, private clinic]

Theme 6: Adapting Techniques and Technology in Digital Therapy Sessions

The sixth theme is placed at the concrete level of therapeutic practice, covering pragmatic adaptations of therapy interventions and technical aspects of the technologies used for VC. One clinician always wore headphones during VC to optimize the dynamics of the conversation, rather than using the speakers:

> *I get the sound directly in the ears instead of it coming from speakers. That makes a difference.* [E3, specialized psychologist, private clinic]

Another clinician talked about how basic VC therapy skills provide great opportunities to carry out therapeutic interventions that traditionally depend on the physical environment of the therapy room. This clinician emphasized the importance of careful verbal guidance and attentive assessment of the patient's needs during VC:

> *I don't feel (that the video format) is an obstacle because I can guide them and see if there's something they need, find a slightly better chair to sit on, or if they have something they can hold in their hand, or if they're working on regulating themselves a bit. Grab a pillow or something like that. They can do that at home as well.* [D3, specialized psychologist using body-oriented methods, private clinic]

Several participants shared the opinion that there is potential for improvement for future VC practices. Some said that simply improving the information about VC directed at patients could strengthen implementation. Others talked about potential improvements in the technology; for example, how AI might enhance the liveliness of VC. Finally, multiple participants talked about the lack of safe and clinically validated software that can mimic whiteboards and paper-based tools that are typically used in traditional assessment and therapy sessions. Such technology already exists and is integrated into VC systems used in other professional contexts, but is not available for clinical use in mental health care.

> *If we get an AI as a video chat function, then we could have more visual dynamics. And I think these dynamics will elevate the experience.* [B2, resident doctor in psychiatry, public hospital]

> *If there were a function where you could incorporate questionnaires or some kind of interactive board in the video call (...) That could perhaps also address some challenges related to forms and documentation.* [A8, specialized psychologist, public hospital]

Discussion

Principal Results

This study explores experiences, attitudes, and current debates related to the role of VC in mental health care and challenges preconceptions identified in our own research environment and beyond. To our knowledge, this is the first qualitative study to explore the experiences and attitudes of mental health professionals across public and private settings on the topic of VC. Moreover, the study included perspectives from a broad selection of services, ranging from those targeting bipolar and psychotic disorders to on-demand services targeting clients with milder conditions and not requiring referrals. Our study identifies differences in VC implementation, experiences, and professional attitudes that support the existing literature

[31,32,42,43]. Contrary to the authors' expectations, key characteristics such as public and private sector affiliation and symptom severity of the target patient populations did not seem to shape the participants' views to a significant degree. Building on this, our study points out important areas to direct future efforts to increase the adoption and impact of VC.

Primarily, our findings demonstrate that higher-level aspects, such as societal values and cultural perceptions, are important in shaping professional attitudes. Psychotherapy and the therapist role are defined through the cultural value systems in society and in professional environments [70]. Thus, fundamental changes in mental health care services, such as the digitalization of the delivery of therapy, can be experienced as a cultural deviation from traditional models of psychotherapy [71,72]. In accordance with this notion, several participants talked about how traditional ideas about psychotherapy weaken the credibility of VC, although many forms of VC-delivered psychotherapy are supported by robust evidence [5-9]. This tendency indicates a lack of cultural familiarity with digitalized treatment and a lack of knowledge about the evidence-based applications of VC.

While the status of VC in clinical environments remains unclear, polarized approaches have been developed. In our study, the participants with the most critical attitudes typically conceptualized VC as part of purely digital, fast-paced, cost-reduction–focused services and, thus, a symptom of an unwanted development in health care. Research on the resistance to digitalized mental health services has identified similar viewpoints of clinicians who experience that technology conflicts with fundamental humanistic values in mental health care, resulting in resistance and stigma related to the digitalization of therapy [72]. This approach facilitates an all-or-nothing view of digitalized services instead of a nuanced perception of VC as a flexible tool that can be integrated in various ways in hybrid services, a model that is becoming increasingly applied in the literature [25-27].

Interestingly, even among the participants who held the most positive attitudes, we found conflicting views on whether VC is feasible for patients with severe conditions. These clinicians described the patient group as vulnerable and conveyed that it was their professional responsibility to protect these patients accordingly. However, adjustments to the service alternatives based on the vulnerable status of patient groups may not always result in better outcomes for the patients [73]. Furthermore, the exclusion of such patient populations may be based on stigmatic perceptions of the patients, which have been found to be relatively frequent among clinicians working in mental health care services [74,75]. Differentiating the treatment offered based on patients' diagnoses may preserve such preconceptions. Although the existing evidence pertaining to VC in services targeting these populations is still scarce and calls for more research [15-19], we encourage an inclusive and personalized approach and advocate refraining from excluding patients based on diagnosis alone.

Importantly, the education of mental health professionals plays a fundamental role in defining the scope of psychotherapy and conceptualizing VC. Currently, VC is mostly absent from the training of mental health personnel, which may promote more trust in traditional cultural perceptions of therapy and the limits of VC. Moreover, mental health professionals report low self-efficacy and a lack of training and support in clinics [14,35,36,47,48]. At the same time, the research field continues to produce new knowledge about VC, such as through studies focused on adapting basic therapeutic skills to the video format and studies on how we can maintain a feeling of connectedness in video sessions [10,14]. In addition, the body of literature covers adaptations of interventions within therapy orientations that are often closely associated with traditional settings and that emphasize sensory experiences and the atmosphere of the therapy room, such as psychodynamic transference [11], emotion-focused therapy exercises [12], and EMDR [13]. Our findings encourage the inclusion of knowledge about VC in the basic training of professionals to ensure that VC is incorporated from an early point.

We also found that the amount of personal experience with VC plays an important role in shaping attitudes, as it directly challenges established cultural assumptions. Several participants expressing optimistic attitudes stated that they had not been able to perceive VC as feasible and valuable before they had practiced it themselves and adjusted their preconceptions. These observations partly align with previous research showing that more experience with VC is associated with positive attitudes and less self-doubt as a therapist [49-51]. These findings are supported by research suggesting that one of the main drivers of adoption is clinicians' own personal experiences, proving that digital tools can result in clinical benefits [76]. Together with the literature, our findings emphasize the importance of personal experience to perceiving the value of VC. Thus, we suggest that services become aware of this tendency, demonstrate the beneficial effects for both service users and providers, and develop incentives for clinicians to integrate VC in their practice.

Finally, our study presents a range of creative and innovative therapy practices and service designs, suggesting that there is potential synergistic value in sharing local insights between professional environments. Several of the service models identified have been implemented only to a very limited degree across the clinical sites of the study, such as the digital hub or digital first sessions (see Theme 2 in the Results section). Some of these service designs have been mentioned in the literature but have not been tested in public settings in Norway [27]. Using insights and service models from both the private sector and innovative public sites may be valuable in future decision-making regarding service designs in public health care. Interestingly, few innovative activities revolve around the technologies used for VC, and clinicians report that the platforms they use lack the basic tools used in traditional therapy [24,41,46]. This indicates that there is room for improvement and further innovation in terms of clinical VC software, which has barely changed since the introduction of VC in mental health care.

Limitations

Although the method chosen for this study does not aim to produce generalizable findings, and although our findings can still be transferable, we wish to explicitly acknowledge that the composition of our sample may have moved the analytical focus and the results in a more optimistic and progressive direction. For instance, several participants recruited from one of the public hospitals were involved in research and development, while two were involved in digitalization projects. We are also aware that the participants who decided to participate might have had more positive attitudes toward the topic of the research project than those who did not want to participate. This, combined with the authors' optimistic attitudes and roles in other digitalization projects closely related to the topic of this study, possibly influenced the angle we chose for this paper, which encourages digitalization in mental health care. Another limitation relates to the practical transferability of findings produced in this study in an international context, as all participants were recruited from Norwegian health care sites. Furthermore, the analysis is not exhaustive in discussing aspects that have an impact on the participants' viewpoints. For instance, regarding the private actors' contextual characteristics, the paper discusses the impact of incentives for practicing VC and whether VC use is voluntary or not to a very limited degree. An important research-related issue resulting from such differences in public and private sector contexts revolved around the recruitment for this study. Participants from private clinics received economic compensation of about US $110 because of slow recruitment due to the nature of the private practice financial model. Lost activity with paying clients results in reduced income, which is a barrier related to participation in research projects. The other participants did not receive similar economic compensation since they were all employed at university hospitals, where participation in research and development projects is integral. Future research should explore the effects of different organizational conditions in detail, which may be of great importance in developing attractive workplaces for mental health professionals and sustainable hybrid services.

Conclusions

This qualitative exploration emphasizes the multidimensional nature of professional perspectives on VC. Our findings underline how overarching societal and cultural ideas influence professional attitudes, logic, and, thus, clinical practice. Our ambition is that the paper challenges preconceptions about digitalized therapy and that it points to useful strategies for targeting problems concerning the limited adoption of VC. Educational institutions and services must promote a nuanced, evidence-based concept of VC and demonstrate benefits for both service users and providers. In addition, incentives for mental health professionals to practice VC should be developed and presented. There is unrealized potential in applying existing therapy adaptations and hybrid service models established in separate parts of health care systems, particularly for public clinics. In addition, there is room for innovation with regard to service designs that integrate VC and the technology platforms used to create unique value in hybrid services and to increase the attractiveness of the VC format in mental health care.

Acknowledgments

We wish to thank all the study participants and the two user representatives for their contributions to this study. This study is funded by the Research Council of Norway through grant 309264 (Forhelse Research Centre for Digital Health Services, a centre for research-based innovation).

Data AVailability

To protect the anonymity of the participants, the qualitative dataset generated during this study will not be available to readers. More information concerning the study settings is available from the corresponding author on reasonable request.

Authors' Contributions

MS wrote the manuscript. The other authors reviewed, modified, and suggested changes during the writing process. KLR, SRA, LNS, KJK, and MS designed and planned the study. KLR, SRA, HMI, and MS recruited participants. MS was the main interviewer. KLR participated as a supervisor and cointerviewer in the first 3 interviews. MS led the analysis with guidance from the other authors. KLR is the main supervisor of MS as a PhD candidate. SRA and LNS are cosupervisors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures.
[DOCX File , 80 KB - humanfactors_v13i1e80812_app1.docx]

Multimedia Appendix 2
Interview guide.
[DOCX File , 17 KB - humanfactors_v13i1e80812_app2.docx]

Multimedia Appendix 3

Detailed description of the analysis.
[DOCX File , 18 KB - humanfactors_v13i1e80812_app3.docx]

References

1. Picillo B, Yu-Lefler H, Bui C, Wendt M, Sripipatana A. Telehealth-facilitated mental health care access and continuity for patients served at the Health Resources and Services Administration-funded health centers. Telemed J E Health 2025;31(7):838-847. [doi: 10.1089/tmj.2025.0011] [Medline: 40014364]
2. Marini C, Steinberg MH, Miron CD, Irwin M, Schantz BL, Ginsberg DL, et al. Examination of patient and provider satisfaction, benefits, and challenges with psychiatric outpatient and hospital-based telehealth treatment during the COVID-19 pandemic. Psychiatr Q 2025;96(4):869-885. [doi: 10.1007/s11126-025-10150-w] [Medline: 40268852]
3. Cummings JR, Kalk T, Trello S, Walker ER, Graetz I. Therapists' perspectives on access to telemental health among medicaid-enrolled youth. Am J Manag Care 2023;29(11):e339-e347 [FREE Full text] [doi: 10.37765/ajmc.2023.89430] [Medline: 37948654]
4. Galvin E, Desselle S, Gavin B, Quigley E, Flear M, Kilbride K, et al. Patient and provider perspectives of the implementation of remote consultations for community-dwelling people with mental health conditions: a systematic mixed studies review. J Psychiatr Res 2022;156:668-678 [FREE Full text] [doi: 10.1016/j.jpsychires.2022.10.051] [Medline: 36399859]
5. Hagi K, Kurokawa S, Takamiya A, Fujikawa M, Kinoshita S, Iizuka M, et al. Telepsychiatry versus face-to-face treatment: systematic review and meta-analysis of randomised controlled trials. Br J Psychiatry 2023;223(3):407-414 [FREE Full text] [doi: 10.1192/bjp.2023.86] [Medline: 37655816]
6. Oh J, Ho J, Lee S, Park J. Effects of digital psychotherapy on suicide: a systematic review and meta-analysis. Healthcare (Basel) 2024;12(14) [FREE Full text] [doi: 10.3390/healthcare12141435] [Medline: 39057578]
7. Shaker AA, Austin SF, Storebø OJ, Schaug JP, Ayad A, Sørensen JA, et al. Psychiatric treatment conducted via telemedicine versus in-person modality in posttraumatic stress disorder, mood disorders, and anxiety disorders: systematic review and meta-analysis. JMIR Ment Health 2023;10:e44790 [FREE Full text] [doi: 10.2196/44790] [Medline: 37277113]
8. Giovanetti AK, Punt SEW, Nelson E, Ilardi SS. Teletherapy versus in-person psychotherapy for depression: a meta-analysis of randomized controlled trials. Telemed J E Health 2022;28(8):1077-1089. [doi: 10.1089/tmj.2021.0294] [Medline: 35007437]
9. Krzyzaniak N, Greenwood H, Scott AM, Peiris R, Cardona M, Clark J, et al. The effectiveness of telehealth versus face-to face interventions for anxiety disorders: a systematic review and meta-analysis. J Telemed Telecare 2024;30(2):250-261. [doi: 10.1177/1357633X211053738] [Medline: 34860613]
10. Kumar MS, Krishnamurthy S, Dhruve N, Somashekar B, Gowda MR. Telepsychiatry netiquette: connect, communicate and consult. Indian J Psychol Med 2020;42(5 Suppl):22S-26S [FREE Full text] [doi: 10.1177/0253717620958170] [Medline: 33354059]
11. Sayers J. Online Psychotherapy: transference and countertransference issues. Brit J Psychotherapy 2021;37(2):223-233. [doi: 10.1111/bjp.12624]
12. Banack KD. Adapting emotion-focused therapy for teletherapy. Pers Centered Exp Psychother 2021;20(4):303-311. [doi: 10.1080/14779757.2021.1993968]
13. Kaptan SK, Kaya ZM, Akan A. Addressing mental health need after COVID-19: a systematic review of remote EMDR therapy studies as an emerging option. Front Psychiatry 2023;14:1336569 [FREE Full text] [doi: 10.3389/fpsyt.2023.1336569] [Medline: 38250261]
14. Werbart A, Dominicus E, Sheahan L, Philips B. When the screen deepens the sense of connectedness: what therapists have found effective in successful cases of teletherapy. Couns and Psychother Res 2024;25(1). [doi: 10.1002/capr.12820]
15. Farrell A, George N, Amado S, Wozniak J. A systematic review of the literature on telepsychiatry for bipolar disorder. Brain Behav 2022;12(10):e2743 [FREE Full text] [doi: 10.1002/brb3.2743] [Medline: 36102239]
16. Greenwood H, Krzyzaniak N, Peiris R, Clark J, Scott AM, Cardona M, et al. Telehealth versus face-to-face psychotherapy for less common mental health conditions: systematic review and meta-analysis of randomized controlled trials. JMIR Ment Health 2022;9(3):e31780 [FREE Full text] [doi: 10.2196/31780] [Medline: 35275081]
17. Randall L, Raisin C, Waters F, Williams C, Shymko G, Davis D. Implementing telepsychiatry in an early psychosis service during COVID-19: experiences of young people and clinicians and changes in service utilization. Early Interv Psychiatry 2023;17(5):470-477 [FREE Full text] [doi: 10.1111/eip.13342] [Medline: 35943177]
18. Lal S, Abdel-Baki A, Lee H. Telepsychiatry services during COVID-19: a cross-sectional survey on the experiences and perspectives of young adults with first-episode psychosis. Early Interv Psychiatry 2023;17(4):368-377 [FREE Full text] [doi: 10.1111/eip.13332] [Medline: 35731034]
19. Best MW, Romanowska S, Zhou Y, Wang L, Leibovitz T, Onno KA, et al. Efficacy of remotely delivered evidence-based psychosocial treatments for schizophrenia-spectrum disorders: a series of systematic reviews and meta-analyses. Schizophr Bull 2023;49(4):973-986 [FREE Full text] [doi: 10.1093/schbul/sbac209] [Medline: 37066769]
20. Steidtmann D, McBride S, Mishkind M, Shore J. From couch to click: opportunities and challenges of hybrid work models in mental health care. Curr Psychiatry Rep 2025;27(6):385-392. [doi: 10.1007/s11920-025-01607-7] [Medline: 40186706]

21. Yellowlees P, Richard Chan S, Burke Parish M. The hybrid doctor-patient relationship in the age of technology - telepsychiatry consultations and the use of virtual space. Int Rev Psychiatry 2015;27(6):476-489. [doi: 10.3109/09540261.2015.1082987] [Medline: 26493089]
22. Aafjes-van Doorn K, Békés V, Stukenberg KW, Roberts KE. Therapists’ views and recommendations on the ins and outs of practicing teletherapy. Couns Psychol Q 2023;37(3):398-418. [doi: 10.1080/09515070.2023.2227133]
23. Honey A, Hines M, Barton R, Berry B, Gilroy J, Glover H, et al. Preferences for telehealth: a qualitative study with people accessing a new mental health service. Digit Health 2023;9:20552076231211083 [FREE Full text] [doi: 10.1177/20552076231211083] [Medline: 37928331]
24. Leuchtenberg S, Gromer D, Käthner I. Videoconferencing versus face-to-face psychotherapy: insights from patients and psychotherapists about comparability of therapeutic alliance, empathy and treatment characteristics. Couns Psychother Res 2022;23(2):389-403. [doi: 10.1002/capr.12538]
25. Chen K, Huang JJ, Torous J. Hybrid care in mental health: a framework for understanding care, research, and future opportunities. NPP Digit Psychiatry Neurosci 2024;2(1):16. [doi: 10.1038/s44277-024-00016-7] [Medline: 41360926]
26. Nwokedi CN, Olowe KJ, Alli OI, Iguma DR. The role of telehealth in psychological counseling: a comprehensive review. World J Adv Res Rev 2025;25(1):1639-1649. [doi: 10.30574/wjarr.2025.25.1.0228]
27. Mucić D, Hilty D, Shore J. Telepsychiatry: from "telepathy" to the mainstream. In: Mucić D, Hilty DM, editors. Digital Mental Health: The Future is Now. Cham, Switzerland: Springer Nature Switzerland; 2024:215-254.
28. Hjelle KM, Eide TO, Thorsen AL, Kvale G, Hagen K, Snorrason I, et al. The Bergen 4-day treatment for panic disorder: adapting to COVID-19 restrictions with a hybrid approach of face-to-face and videoconference modalities. BMC Psychiatry 2023;23(1):570 [FREE Full text] [doi: 10.1186/s12888-023-05062-7] [Medline: 37550696]
29. Kassing F, Seeley JR, Rizvi SL, Compton SN, Sinclair J, Oshin LA, et al. Assessing the implementation of suicide-focused treatments delivered in hybrid telemental health format in a real-world setting. Psychother Res 2025;35(7):1198-1212. [doi: 10.1080/10503307.2024.2415104] [Medline: 39439041]
30. Peifer JS, Bradley E, Taasoobshirazi G. Pilot testing a brief partner-inclusive hybrid intervention for perinatal mood and anxiety disorders. Front Psychiatry 2022;13:735582 [FREE Full text] [doi: 10.3389/fpsyt.2022.735582] [Medline: 35633795]
31. Appleton R, Barnett P, Vera San Juan N, Tuudah E, Lyons N, Parker J, et al. Implementation strategies for telemental health: a systematic review. BMC Health Serv Res 2023;23(1):78 [FREE Full text] [doi: 10.1186/s12913-022-08993-1] [Medline: 36694164]
32. Rains LS, Dalton-Locke C, Landau S, Needle JJ, Johnson S. Variations in the uptake of telemental health technologies in community and crisis mental health services during the early pandemic: a survey of mental health professionals in the UK. BMC Psychiatry 2022;22(1):776 [FREE Full text] [doi: 10.1186/s12888-022-04385-1] [Medline: 36494646]
33. Cowan KE, McKean AJ, Gentry MT, Hilty DM. Barriers to use of telepsychiatry: cinicians as gatekeepers. Mayo Clin Proc 2019;94(12):2510-2523. [doi: 10.1016/j.mayocp.2019.04.018] [Medline: 31806104]
34. Connolly SL, Miller CJ, Lindsay JA, Bauer MS. A systematic review of providers' attitudes toward telemental health via videoconferencing. Clin Psychol (New York) 2020;27(2) [FREE Full text] [doi: 10.1111/cpsp.12311] [Medline: 35966216]
35. Negev M, Magal T, Kaphzan H. Attitudes of psychiatrists toward telepsychiatry: a policy Delphi study. Digit Health 2023;9:20552076231177132 [FREE Full text] [doi: 10.1177/20552076231177132] [Medline: 37312951]
36. Haydon HM, Fowler JA, Taylor ML, Smith AC, Caffery LJ. Psychological factors that contribute to the use of video consultations in health care: systematic review. J Med Internet Res 2024;26:e54636 [FREE Full text] [doi: 10.2196/54636] [Medline: 39661977]
37. Moeller AM, Christensen LF, Hansen JP, Andersen PT. Patients' acceptance of video consultations in the mental health services: a systematic review and synthesis of qualitative research. Digit Health 2022;8:20552076221075148 [FREE Full text] [doi: 10.1177/20552076221075148] [Medline: 35154803]
38. Almathami HKY, Win KT, Vlahu-Gjorgievska E. Barriers and facilitators that influence telemedicine-based, real-time, online consultation at patients' homes: systematic literature review. J Med Internet Res 2020;22(2):e16407 [FREE Full text] [doi: 10.2196/16407] [Medline: 32130131]
39. Samad L, Teague B, Elzubeir K, Moreira K, Agarwal N, Bagge S, et al. Mental health care using video during COVID-19: service user and clinician experiences, including future preferences. MHRJ 2023;28(1):93-111. [doi: 10.1108/MHRJ-06-2022-0036]
40. Guinart D, Marcy P, Hauser M, Dwyer M, Kane JM. Mental health care providers' attitudes toward telepsychiatry: a systemwide, multisite survey during the COVID-19 pandemic. Psychiatr Serv 2021;72(6):704-707. [doi: 10.1176/appi.ps.202000441] [Medline: 33593104]
41. Buckman JEJ, Saunders R, Leibowitz J, Minton R. The barriers, benefits and training needs of clinicians delivering psychological therapy via video. Behav Cogn Psychother 2021;49(6):696-720. [doi: 10.1017/S1352465821000187] [Medline: 33966666]
42. Roddy JK, Gabriel L, Sheehy R, Charura D, Dunn E, Hall J, et al. Mental health practitioners’ perceptions of online working: a literature review. Br J Guid Couns 2024;53(2):190-205. [doi: 10.1080/03069885.2024.2373179]

43. Békés V, Aafjes-van Doorn K, Roberts KE, Stukenberg K, Prout T, Hoffman L. Adjusting to a new reality: Consensual qualitative research on therapists' experiences with teletherapy. J Clin Psychol 2023;79(5):1293-1313. [doi: 10.1002/jclp.23477] [Medline: 36704974]
44. Aafjes-Van Doorn K, Békés V, Luo X, Hopwood CJ. Therapists' perception of the working alliance, real relationship and therapeutic presence in in-person therapy versus tele-therapy. Psychother Res 2024;34(5):574-588. [doi: 10.1080/10503307.2023.2193299] [Medline: 37011405]
45. Gullslett MK, Kristiansen E, Nilsen ER. Therapists' experience of video consultation in specialized mental health services during the COVID-19 pandemic: qualitative interview study. JMIR Hum Factors 2021;8(3):e23150 [FREE Full text] [doi: 10.2196/23150] [Medline: 34096505]
46. Åhs JW, Ranheim A, Eriksson H, Mazaheri M. Encountering suffering in digital care: a qualitative study of providers' experiences in telemental health care. BMC Health Serv Res 2023;23(1):418 [FREE Full text] [doi: 10.1186/s12913-023-09367-x] [Medline: 37127655]
47. Lipschitz JM, Connolly SL, Van Boxtel R, Potter JR, Nixon N, Bidargaddi N. Provider perspectives on telemental health implementation: lessons learned during the COVID-19 pandemic and paths forward. Psychol Serv 2023;20(Suppl 2):11-19 [FREE Full text] [doi: 10.1037/ser0000625] [Medline: 35201809]
48. Sharma G, Devan K. The effectiveness of telepsychiatry: thematic review. BJPsych Bull 2023;47(2):82-89 [FREE Full text] [doi: 10.1192/bjb.2021.115] [Medline: 34915955]
49. Burbach FR, Stiles KM, Brown K, Allen VME, Mc Kenny R. 'More collaborative and open': as systemic psychotherapists become more familiar with videoconferencing, are the advantages of the medium becoming more apparent? A focus group study with UK systemic practitioners. Hum Syst 2022;2(2):82-101. [doi: 10.1177/26344041221085227]
50. Aafjes-van Doorn K, Békés V, Prout TA. Grappling with our therapeutic relationship and professional self-doubt during COVID-19: will we use video therapy again? Counsel Psychol Q 2020;34(3-4):473-484. [doi: 10.1080/09515070.2020.1773404]
51. Kaphzan H, Sarfati Noiman M, Negev M. The attitudes and perceptions of israeli psychiatrists toward telepsychiatry and their behavioral intention to use telepsychiatry. Front Psychiatry 2022;13:829965 [FREE Full text] [doi: 10.3389/fpsyt.2022.829965] [Medline: 35386519]
52. Psykiske helsetjenester og tilgang til psykologtjenester i Norge En analyse av tilgang til og fordeling av helsetjenester for psykiske lidelser og psykologenes rolle i tjenestetilbudet. Oslo Economics. 2022. URL: https://osloeconomics.no/wp-content/uploads/2022/06/Psykiske-helsetjenester-og-tilgang-til-psykologtjenester-i-Norge-26.09.22.pdf [accessed 2025-05-01]
53. Helseforsikring. Finans Norge. 2024. URL: https://www.finansnorge.no/tema/statistikk-og-analyse/forsikring/helseforsikring/ [accessed 2025-05-01]
54. Achtyes ED, Glenn T, Monteith S, Geddes JR, Whybrow PC, Martini J, et al. Telepsychiatry in an era of digital mental health startups. Curr Psychiatry Rep 2023;25(6):263-272 [FREE Full text] [doi: 10.1007/s11920-023-01425-9] [Medline: 37166622]
55. Larsson A, Elf O, Gross C, Elf J. Welfare services in an era of digital disruption: how digitalization reshapes the health care market. In: Digital Transformation and Public Services. Milton Park, Oxfordshire, UK: Routledge; 2019:33-57.
56. Frank HE, Cruden G, Crane ME. Private practice, private insurance, and private pay mental health services: an understudied area in implementation science. Adm Policy Ment Health 2024;51(1):1-6 [FREE Full text] [doi: 10.1007/s10488-023-01306-6] [Medline: 37880471]
57. Mohler MJ. Collaboration across clinical silos. Front Health Serv Manage 2013;29(4):36-44. [Medline: 23858987]
58. Pedersen ERG, Sudzina F, Rosati F. A multi-dimensional study of organisational boundaries and silos in the healthcare sector. Health Serv Manage Res 2024;37(4):200-208. [doi: 10.1177/09514848231218617] [Medline: 38011078]
59. Snell T. Breaking down silos: the value of interdisciplinary production of knowledge for (health) innovation. In: Handbook of Ethnography in Healthcare Research. Milton Park, Oxfordshire, UK: Routledge; 2020:388-400.
60. Ii SS, Fitzgerald L, Morys-Carter MM, Davie NL, Barker R. Knowledge translation in tri-sectoral collaborations: an exploration of perceptions of academia, industry and healthcare collaborations in innovation adoption. Health Policy 2018;122(2):175-183. [doi: 10.1016/j.healthpol.2017.11.010] [Medline: 29191387]
61. Lander B. Sectoral collaboration in biomedical research and development. Scientometrics 2012;94(1):343-357. [doi: 10.1007/s11192-012-0776-8]
62. Robinson OC. Sampling in interview-based qualitative research: a theoretical and practical guide. Qual Res Psychol 2013;11(1):25-41. [doi: 10.1080/14780887.2013.801543]
63. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2020;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
64. Saunes I, Durvy B, Litvinova Y. Norway: health system summary. European Observatory on Health Systems and Policies, WHO Regional Office for Europe. 2024. URL: https://eurohealthobservatory.who.int/publications/i/norway-health-system-summary-2024 [accessed 2026-01-08]
65. The basic structure of the Norwegian health service. Ministry of Health and Care Services. 2024. URL: https://www.regjeringen.no/en/topics/health-and-care/innsikt/the-basic-structure-of-the-norwegian-health-service/id227440/ [accessed 2025-05-01]

66. Slevin E, Sines D. Enhancing the truthfulness, consistency and transferability of a qualitative study: utilising a manifold of approaches. Nurse Researcher 2000;7(2):79-98. [doi: 10.7748/nr2000.01.7.2.79.c6113]
67. Skoge M, Aminoff SR, Barrett EA, Bryhni GE, Kling K, Kværner KJ, et al. A mobile app designed to promote shared decision-making in the treatment of psychotic disorders: feasibility and acceptability study. JMIR Hum Factors 2025;12:e68813 [FREE Full text] [doi: 10.2196/68813] [Medline: 40644621]
68. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qual Psychol 2022;9(1):3-26. [doi: 10.1037/qup0000196]
69. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
70. Tseng WS. Culture and psychotherapy: review and practical guidelines. Transcult Psychiatry 1999;36(2):131-179. [doi: 10.1177/136346159903600201]
71. Docherty JP, Colbert BM. The evolution of psychotherapy: from Freud to prescription digital therapeutics. Front Psychiatry 2024;15:1477543 [FREE Full text] [doi: 10.3389/fpsyt.2024.1477543] [Medline: 39507283]
72. Berardi C, Antonini M, Jordan Z, Wechtler H, Paolucci F, Hinwood M. Barriers and facilitators to the implementation of digital technologies in mental health systems: a qualitative systematic review to inform a policy framework. BMC Health Serv Res 2024;24(1):243 [FREE Full text] [doi: 10.1186/s12913-023-10536-1] [Medline: 38408938]
73. Atkinson JM. Protecting or empowering the vulnerable? mental illness, communication and the research process. Res Ethics 2007;3(4):134-138. [doi: 10.1177/174701610700300410]
74. Sullivan G, Mittal D, Reaves CM, Haynes TF, Han X, Mukherjee S, et al. Influence of schizophrenia diagnosis on providers' practice decisions. J Clin Psychiatry 2015;76(8):1068-74; quiz 1074. [doi: 10.4088/JCP.14m09465] [Medline: 26335084]
75. Wang K, Link BG, Corrigan PW, Davidson L, Flanagan E. Perceived provider stigma as a predictor of mental health service users' internalized stigma and disempowerment. Psychiatry Res 2018;259:526-531 [FREE Full text] [doi: 10.1016/j.psychres.2017.11.036] [Medline: 29156425]
76. Feijt MA, de Kort YA, Bongers IM, IJsselsteijn WA. Perceived drivers and barriers to the adoption of emental health by psychologists: the construction of the levels of adoption of eMental health model. J Med Internet Res 2018;20(4):e153 [FREE Full text] [doi: 10.2196/jmir.9485] [Medline: 29691215]

Abbreviations

EMDR: eye movement desensitization and reprocessing therapy
TSD: Services for Sensitive Data
VC: video consultations

Edited by A Kushniruk; submitted 17.Jul.2025; peer-reviewed by M Pulier, J Muschol; comments to author 04.Dec.2025; revised version received 11.Dec.2025; accepted 02.Jan.2026; published 26.Jan.2026.

Please cite as:
Skoge M, Aminoff SR, Ihler HM, Kværner KJ, Støme LN, Romm KL
Exploring Video Consultations Across the Public and Private Sectors in Norway: Semistructured Interview Study
JMIR Hum Factors 2026;13:e80812
URL: https://humanfactors.jmir.org/2026/1/e80812
doi:10.2196/80812
PMID:

Original Paper

Perspectives of Frontline Clinicians and End-Line Users on Smartphone-Based Photography for Assessing Traumatic Dental Injuries: Focus Group Interview Study and Thematic Analysis

Emily C Schultz[1,2], RDH, MS; Boyen Huang[1], DDS, MHA, PhD; Margaret Shenouda[1], BA; Mohamed Estai[3], PhD; Sarbin Ranjitkar[4], BDS, BScDent (Hons), PhD; Jeffrey P Louie[5], MD; Patimaporn Pungchanchaikul[6], DDS, MSc, PhD

[1]Department of Primary Dental Care, University of Minnesota School of Dentistry, Minneapolis, MN, United States
[2]Department of Dental Education, College of Allied Health and Nursing, Minnesota State University, Mankato, MN, United States
[3]School of Human Sciences, The University of Western Australia, Perth, Australia
[4]School of Dentistry, College of Health, Adelaide University, Adelaide, Australia
[5]Department of Pediatrics, University of Minnesota Medical School, Minneapolis, MN, United States
[6]Faculty of Dentistry, Khon Kaen University, Khon Kaen, Thailand

Corresponding Author:
Boyen Huang, DDS, MHA, PhD
Department of Primary Dental Care
University of Minnesota School of Dentistry
515 Delaware Street SE
Minneapolis, MN, 55455
United States
Phone: 1 612 624 7848
Email: huan2321@umn.edu

Abstract

Background: Mobile health (mHealth) is increasingly used in teledentistry for telediagnosis and other services; yet, the perceptions of frontline clinicians and end-line users regarding these technologies remain unexplored.

Objective: This study examined the acceptability, feasibility, and usability of an mHealth model for telediagnosis from the perspectives of frontline clinicians and end-line users.

Methods: A qualitative study using focus group interviews was conducted with 15 participants, including frontline clinicians and end-line users. Frontline clinicians captured dental images via a smartphone app, while end-line users assessed them through an mHealth platform. Interview transcriptions were thematically analyzed using consensus coding.

Results: Thematic analysis identified 9 key themes: feasibility and perceived ease of use, perceived usefulness, compatibility, self-image and social influences, self-efficacy, voluntariness and behavior intention, anxiety, facilitating conditions, and attitudes toward a behavior. Participants considered smartphone-based photography acceptable and feasible for remote dental assessment. Facilitators and barriers to implementing the mHealth model were highlighted, and recommendations for improvements were proposed.

Conclusions: Cyclical education and professional development are essential to enhancing user confidence and technology usability. Addressing patient and clinician resistance through targeted education, improved communication, and operational upgrades such as camera grids, system integration, and simplified login can support adoption. This study highlights mHealth's potential in emergency dental assessment and screening, particularly for underserved populations, and underscores opportunities for interprofessional collaboration. Future research should explore broader clinical applications across oral health conditions.

(JMIR Hum Factors 2026;13:e82668) doi:10.2196/82668

KEYWORDS

acceptability; dental trauma; feasibility; image quality; mHealth; teledentistry; telediagnosis; usability; user perception

Introduction

Traumatic dental injuries (TDIs) are prevalent [1] and demand urgent attention when accompanied by pain and/or bleeding [2]. Most patients with TDI presenting in emergency departments need consultation and evaluation from dental professionals [2]. However, the limited on-call availability of dental professionals in emergency settings often leads to unnecessary patient transfers and increased health care costs [3], an issue that worsened as dental emergency services declined during the COVID-19 pandemic [4]. To address these challenges, technologies for virtual dental examinations have been developed and implemented [5].

Mobile health (mHealth) facilitators have incorporated mobile phones and wireless technological devices into health care practices, supporting the promotion and maintenance of health, enhancing preventive care, improving clinical decision-making and operational efficiency, and enabling remote communication and interaction. This approach has gained widespread adoption among health care professionals and caregivers [6]. Key benefits of mHealth include increased accessibility to health care services, the provision of anonymous consultation, and decreased reliance on travel and physical contact. These advantages were particularly valuable during the COVID-19 pandemic, when social distancing measures were necessary [7].

The concept of mHealth as a modality of teledentistry, or more specifically, m-oral health [8], includes teleconsultation for treatment planning and review [9], telediagnosis with virtual examination [9-14], telemonitoring using patient-generated health data [15-18], telesupport with interactive multimedia [19], and teleintervention to improve therapy adherence and complication management [20]. Prior mHealth studies have demonstrated adequate diagnostic performance for the remote assessment of TDIs [10], dental caries [11], and oral cancer [12] using smartphone-acquired photographs. In a survey conducted in Saudi Arabia, more than 60% of dentists have used mHealth technologies to capture and/or transmit clinical photographs, and the majority of them were confident in the diagnostic accuracy using this approach [21].

Our recent work described a workflow for telediagnosis of TDI, where trained clinicians and students used a smartphone camera and app to capture dental photographs, uploading them to secure cloud storage. Remote dentists and dental therapists then reviewed the images, recorded dental findings, and provided clinical recommendations through the same pathway in reverse. While the quantitative performance metrics of this mHealth practice were promising, areas for improvement, such as image quality and professional development on dental photography and remote assessment, were also indicated [10].

Within the Technology Acceptance Model (TAM), perceived ease of use and perceived usefulness are the determinants for an individual’s intention to use a new technology [22]. Based on different roles and concerns, Wallis et al [23] categorized mHealth users as patients, frontline clinicians (point-of-care clinicians), and end-line users (academic experts). Building on a similar classification, the users of our telediagnosis model comprised human participants (patients and nonpatient individuals who provided consent for dental photography), frontline clinicians (those who captured dental photos), and end-line users (those who remotely assessed dental photos) [10]. When searching literature with keywords of feasibility, acceptability, usability, or user perception, we found that only a few studies have reported frontline clinicians’ and end-line users’ perspectives on specific mHealth models, practices, or technologies for dental care or oral health promotion [13,15-18,20]. Among these, only 1 study focused on telediagnosis [13], but their diagnostic method was not photography-related. Of further note, an earlier mHealth study reported the perspectives of parents and caregivers on a photographic telediagnosis model for the detection of dental caries [14], but they did not investigate the perspectives of frontline clinicians and end-line users.

Despite evidence that smartphone-acquired photographs can achieve acceptable diagnostic performance for several oral conditions, few studies have examined the perspectives of the 2 professional user groups essential to photographic telediagnosis, frontline clinicians who capture images and end-line users who interpret them. Prior work has focused on patients, caregivers, or nonphotography telemethods, leaving clinician-centered implementation questions unanswered. This gap matters because diagnostic accuracy alone does not guarantee real-world adoption. Human factors such as privacy concerns, workflow fit, training needs, and system integration also directly influence acceptability and sustained use. By exploring these clinician perspectives, this study addresses practical barriers and facilitators that must be resolved to translate promising diagnostic performance into scalable, safe, and equitable teledentistry services. This study therefore aimed to explore frontline clinicians’ and end-line users’ perspectives on the feasibility, usability, and acceptability of smartphone-based photographic telediagnosis for TDI.

Methods

Study Design and Setting

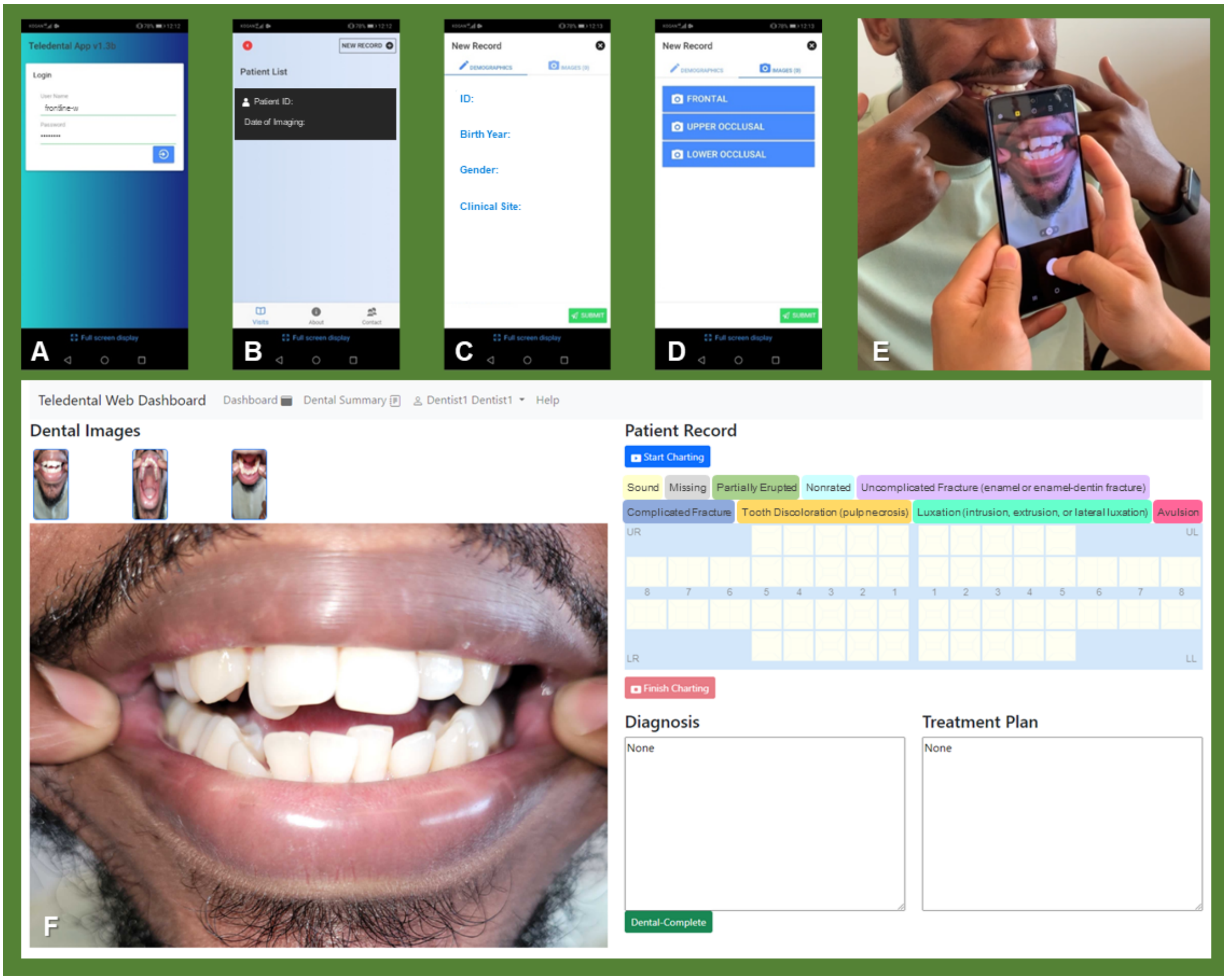
This qualitative study was part of a larger project and was reviewed and approved by the University of Minnesota Institutional Review Board (Study ID: STUDY00014736). The study setting consisted of multiple sites located in Minnesota (United States) and Khon Kaen (Thailand), including the University of Minnesota clinical site, Khon Kaen University clinical sites, and Minnesota State Fair research facilities. The quantitative component of the larger project has been reported recently [10]. Following that, a qualitative study design grounded in the TAM [22] and interpretivism [24] was used to evaluate the perspectives of frontline clinicians (dental photographers) and end-line users (remote reviewers) on their use of the mHealth technology to assess TDI. Qualitative study data were collected in November and December 2023 and analyzed from January to August 2024. The conduction and presentation of this study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [25].

This study built upon our quantitative work of TDI remote assessments using an image acquisition app “Teledental” (Commonwealth Scientific and Industrial Research Organisation

[CSIRO]) and a web-based data management platform “Remote-I” (CSIRO; Figure 1). From August 2022 to July 2023, the photographs of teeth were captured by 11 trained frontline clinicians using the Teledental app, and then remotely reviewed by 5 licensed dental professionals (end-line users) using the Remote-I system [10].

In Figure 1, panels A-C show the “Teledental” app interface used by frontline clinicians to log in with individual passwords and enter patient data, including patient ID, birth year, sex, and clinical site. Panel D displays the app’s imaging screen, which prompts users to capture 3 standardized intraoral views: frontal, upper occlusal, and lower occlusal. Panel E depicts 2 frontline clinicians participating in one-on-one calibration after completing a training session designed to ensure consistent image quality and adherence to capture protocols. Panel F shows the “Remote-I” platform interface used by end-line users to access the patient record created during calibration; when the cursor hovers over a thumbnail, the corresponding dental image is enlarged for review. End-line users examined all 3 images and charted dental conditions using a color-coded odontogram that included diagnostic categories such as uncomplicated fractures, luxations, and partially erupted teeth. The clinicians provided consent for the use of their images and data in the publication.

Figure 1. Workflow and interfaces of the mHealth app and platform used in this study for photographic telediagnosis of traumatic dental injuries.



User Name: frontline-w

Diagnosis: None

Treatment Plan: None

Basic Security Measures

All researchers, frontline clinicians, and end-line users adhered to institutional review board–approved procedures governing data access and privacy. To limit image persistence on devices, the Teledental app was configured to prevent local storage of captured images; images were transmitted directly to a US-owned cloud service selected jointly by the research team and the vendor (CSIRO). Remote image review required authenticated access to the Remote-I platform; the platform’s settings disallowed image downloads to end-line users’ workstations. Study smartphones were purchased new from the vendor (Samsung), configured with restricted user permissions, and connected only to password-protected institutional Wi-Fi (eg, eduroam or hospital internal networks); installation of additional apps was prohibited; devices were centrally stored and managed by a key researcher in Minnesota and a key researcher in Khon Kaen. Individual user accounts and passwords were required for both Teledental and Remote-I and the passwords were changed periodically. Two-factor authentication was not used by study participants during data collection. All study personnel completed training in privacy and data-handling best practices, including HIPAA (Health

Insurance Portability and Accountability Act) principles [26], and compliance with those requirements was monitored.

Participants

All eligible frontline clinicians and end-line users who contributed to the preceding quantitative study were invited (consecutive sampling) and the full available cohort consented to participate, with the only exclusion being a frontline clinician who is a coauthor to avoid conflict of interest. The sample included a purposeful mix of roles and experience levels (dental students, general dentists, dental therapists, and dental specialists) across 2 countries. No participants were selected or excluded by demographic or professional criteria. To acknowledge potential recruitment influences, we report that University of Minnesota employees and student workers received modest incentives (US $40 and US $15, respectively) while participants based in Thailand did not. All participants were notified of the requirement to attend a 60-minute focus group interview through the Zoom app (Zoom Video Communications). Each focus group was composed of 1 to 6 participants according to their schedule availability. The participants based in the United States and Thailand were separated into different focus groups due to different time zones. The interviews were conducted in English by 2 US-based researchers together. All participants were capable of communicating in English during the Zoom interviews.

Data Collection and Instrument

Two moderators, ECS (a female dental hygiene faculty member with a Master of Science degree) and MS (a female Honors college student with a Bachelor of Arts degree) conducted each focus group interview, while BH (a male senior dentistry faculty member and the principal investigator [PI] with a Doctor of Philosophy and Doctor of Dental Surgery degrees) observed all sessions. Prior to the first interview, the moderators underwent training to adopt a neutral interviewing stance, ensuring that participants’ perspectives were accurately captured. Since participants were recruited from frontline and end-line users of a previous study led by the observer, the moderators had a limited prior acquaintance with 3 participants. To maintain transparency, the moderators and observer introduced themselves and outlined the research objectives and content on the informed consent form and at the beginning of each interview. Before the interview with participants based in Thailand, the observer and moderators interacted with them in English for approximately 15 minutes via Zoom to confirm their proficiency and ensure effective communication. Only the participants and researchers were present during the interviews. Field notes were taken by the moderators during the sessions, and all interviews were audio-recorded and securely stored on a cloud server. After each focus group, participants were emailed the full study instrument and given 2 weeks to provide additional comments. To safeguard data quality and reduce risks of misinterpretation, a researcher (PP), a native Thai speaker fluent in English with graduate training in Britain, attended the Thai focus group to provide Zoom technical support and reviewed the transcripts from that session for accuracy. Although data saturation was defined as the point at which no new information emerged across 3 consecutive focus groups [27], we chose to interview all 15 participants in an effort to generate more evidence and decrease bias due to our small but very diverse sample size. Repeat interviews were not conducted.

The study instrument used for the interviews consisted of guided interview questions, exploring the perspectives of the frontline clinicians and end-line users. The interview questions were adapted from published resources [22,28-32] and were rephrased to reflect the context of this qualitative study and the mHealth technologies used. Subsequently, a multimember researcher panel with diverse expertise, including qualitative study design, medical informatics, clinical dentistry, dental public health, and special needs dentistry, reviewed and finalized the questions. The interview questions included in this study are presented in Table 1. The purpose of using this guided interview instrument for the interview was to gain a well-rounded perspective from the participants on their experience with the mHealth model and maintain consistency among all focus group interviews.

Table 1. Guided interview questions adapted from the Technology Acceptance Model.

Interview questions	References
• Tell us about the feasibility of using the mHealth[a] app or platform. What factors played a role?	[22,28]
• How do you feel this mHealth app or platform can contribute to working with dental trauma cases? (In terms of efficiency, usefulness, etc)	[22,28]
• How does this mHealth app or platform fit with the way dental trauma cases are examined and diagnosed?	[28,29]
• How do you feel the use of the mHealth app or platform may change the way you and other people in your profession, specialty or field are viewed?	[28,29]
• Did your previous experiences influence your use of the mHealth app or platform? Anything positive? Anything negative?	[28,30]
• How willing are you to use the mHealth app or platform to examine and diagnose dental trauma cases in the future?	[28,29]
• How likely are you to use an mHealth app or platform like the one we used in the future for dental trauma cases?	[28,31]
• How comfortable or uncomfortable did you feel about using the technology for dental trauma cases? • Only to frontline clinicians: How might you have observed either potential distress or comfort from the patients or fairgoers enrolled in this study?	[28,30]
• What attitudes do you believe your superiors have around your use of the mHealth app or platform to examine and diagnose dental trauma cases?	[28,31]
• What resources, knowledge, or assistance do you feel should be available when using the mHealth app or platform to examine and diagnose dental trauma cases?	[28,32]
• What do you feel went well with the mHealth app or platform? What improvements do you feel could be made with the mHealth app or platform?	[28,31]

[a] mHealth: mobile health.

Data Analysis

A consensus coding process was used to summarize and synthesize the qualitative data from the interview questions and determine the results [33,34]. Coders calibrated via Zoom before conducting any focus group interviews. This early calibration ensured alignment in coding practices and interpretation prior to engaging with the data, to maintain consistency across all transcripts. Memoing was used throughout the coding process to document coder reflections, emerging themes, and the rationale behind coding decisions. These memos supported transparency and helped track the coding development. The research team used a negotiated agreement strategy for intercoder reliability. Prior to data collection, coders collaboratively established a process for resolving discrepancies. When disagreements arose during coding, coders revisited the transcript, discussed their individual interpretations, and reached a consensus. The PI, who observed all focus group interviews, was also consulted to help resolve disagreements and ensure consistency. We used a hybrid coding approach that integrated both deductive and inductive strategies to analyze the focus group data. Initially, we developed a set of a priori codes grounded in our research questions and informed by the TAM [22], which provided a theoretical framework for examining participants' responses. As coding progressed, we remained open to emergent patterns and concepts, allowing for the creation of new inductive codes that captured unanticipated insights arising directly from the data. This flexible approach enabled us to systematically explore expected themes while also incorporating novel findings that enriched our understanding of participants' experiences [34,35]. After all interview sessions had taken place, a Zoom meeting among the research team was held to analyze and discuss the qualitative data, use the transcripts to identify common themes, and extract appropriate evidence, such as descriptive terms and quotes, from the participants' responses. The transcripts and findings were not returned to participants for review.

Ethical Considerations

This study was reviewed and approved by the University of Minnesota Institutional Review Board (Study ID: STUDY00014736). Written informed consent was obtained from all participants prior to data collection. All researchers and study personnel adhered to institutional review board–approved procedures governing data access and privacy. Compensation was provided to University of Minnesota employees and student workers, but not Thailand-based participants due to University out of country payment restrictions.

Results

Overview

A total of 7 interview sessions were conducted. Among the 15 study participants, 10 were frontline clinicians who took the dental photographs and 5 were end-line users who interpreted the dental photographs from a distance. The 10 frontline clinicians included 5 general dentists, 1 pediatric dental specialist, 1 dental therapist, and 3 dental students. The 5 end-line users consisted of 1 pediatric dental specialist, 2 general dentists, and 2 dental therapists. The 6 dentists who acted as frontline clinicians and took dental photographs were all based in Khon Kaen (Thailand), while the other 4 frontline clinicians and the 5 end-line users were based in Minnesota (United States). Table 2 lists the focus groups, roles, professions, sex, and countries of the participants included in this qualitative study.

Data saturation was achieved within the 7 interviews as no new codes or themes emerged after the first 4 interviews. Thematic analysis of frontline clinicians' and end-line users' structured interviews revealed 9 main themes: feasibility and perceived ease of use, perceived usefulness, compatibility, self-image and social influences, self-efficacy, voluntariness and behavior intention, anxiety, facilitating conditions, and attitudes toward a behavior (Table 3).

Table 2. Focus groups, roles, professions, sex, countries, and years in practice of the participants (9 in the United States and 6 in Thailand); years in practice ranged from 0 to 20 years. All consented to publication.

Focus group and role	Profession	Sex	Country	Years in practice
1				
End-line user	Dental therapist	Male	United States	10
End-line user	General dentist	Male	United States	15
2				
End-line user	Dental therapist	Female	United States	2
3				
Frontline clinician	Dental student	Male	United States	0
4				
Frontline clinician	Dental therapist	Female	United States	12
Frontline clinician	Dental student	Male	United States	0
5				
End-line user	General dentist	Male	United States	3
End-line user	Dental specialist	Male	United States	20
6				
Frontline clinician	Dental student	Female	United States	0
7				
Frontline clinician	General dentist	Female	Thailand	5
Frontline clinician	General dentist	Female	Thailand	6
Frontline clinician	General dentist	Female	Thailand	12
Frontline clinician	General dentist	Male	Thailand	5
Frontline clinician	General dentist	Male	Thailand	8
Frontline clinician	Dental specialist	Male	Thailand	15

Table 3. Subthemes (code words), definitions, and representative transcript excerpts that contributed to theme construction.

Main theme	Definition	Subtheme	Example from transcripts
• Feasibility and perceived ease of use	• Statements that describe the feasibility and ease of use of the mHealth[a]	• Consistency • Login • Template	• "I think it was fairly easy for me just because I was comfortable using intraoral cameras before. I think it's something that would be kind of intuitive to a new user. It wasn't that difficult." • "I think the only negative thing, which we kind of figured out, was if the app was left unutilized for a certain amount of time, then it would time out and you'd have to sign back in." • "I want a guideline with the application while using. Kind of like a template to make sure pictures are being taken at correct angles."
• Perceived usefulness	• Comments relating to the usefulness of mHealth	• Access • Timely • Useful	• "I think it's beneficial because, in terms of increasing access, people might not always have access to dental services right away following trauma." • "mHealth could help get the treatment done faster if you could get diagnosis back quicker and in turn get patients treatment faster." • "Specifically speaking about usability, it was really user-friendly in terms of clicking on patient, reviewing what is in odontogram and chart."
• Compatibility	• Statements referring to the familiarity and integration of mHealth technologies	• Technology familiarity	• "I think with our generation in this day and age it is easier to use apps and the app wasn't that complicated at all, it was kind of like straightforward."
• Self-image and social influences	• Comments that describe the social perception of mHealth use	• Acceptance • Resistance	• "I feel like this would be viewed in a positive manner, in that we are taking extra steps to expand access to care through teledentistry and through this mHealth method of like this is easily accessible for patients." • "Some people might say well that's not truly seeing the patient or diagnosing the issue if you're not seeing them in-person since it's kind of a new area of dentistry, but I don't think I'd have any issue with performing teledentistry."
• Self-efficacy	• Statements that describe the user's beliefs in themselves to use mHealth	• Confidence	• "I was very comfortable with how to utilize the technology and then since I've used my phone a lot and taken a lot of images before, it was pretty streamline for me to be able to use."
• Voluntariness and behavior intention	• Comments referring to the user's intention to adopt or integrate mHealth into practice	• Reluctance • Intention	• "For me, I don't feel like it adds a lot to either the patient experience or patient care at the moment." • "I think I would definitely be willing to use it, like I was saying, like for screening patients, I think it's a very valuable asset to see how acute or how soon they would need to get in."

Main theme	Definition	Subtheme	Example from transcripts
- Anxiety	- Statements that describe the user's comfort level related to the use of mHealth	- Comfortable	- "I think it was very easy for me. I think it was something that I felt very comfortable being able to use, taking the images, making sure the images came out well. I felt very comfortable doing that. I know it may be more challenging for people who are not as familiar with being able to use that technology."
- Facilitating conditions	- Statements that describe user's view on what conditions are needed to implement mHealth	- Resources - More practice	- "I think the research team did a good job training us through this, the app, and showing us how to use it, and they even did a one-on-one calibration when we got the phones from them." - "I think if we did a few practice sessions it would help us a lot to be able to emulate what type of images we wanted to take and how we wanted them to look like. So providing some model images of what the appropriate images would look like, I think that would be beneficial."
- Attitudes toward a behavior	- Statements that describe a user's feelings about the use of mHealth	- Positive attitudes - Improvements	- "I was excited when I was using it. I thought it's something that I can definitely make a difference." - "Technology that would streamline the photo taking process, instead of needing to click. Something that would eliminate the need to save or change to the next photo manually. Also not having to login every time, save login for quicker future login."

[a] mHealth: mobile health.

Feasibility and Perceived Ease of Use

As a whole, the frontline clinicians based in Thailand and the United States found the mHealth app feasible and straightforward to use, stating the app itself was self-explanatory, simple, and easy to use. The frontline clinicians at the Minnesota State Fair (United States) used the technology more often and more consistently, while the frontline clinicians at clinical sites in Minnesota (United States) and Khon Kaen (Thailand) used the technology more sporadically. With this being said, frontline clinicians who used it more consistently found it easier to use than those who used it sporadically as they were more familiar with the technology.

Almost all frontline clinicians felt that a more streamlined login process is necessary to improve mHealth app use. They criticized the need to log in each use with a username and password, as it could be time-consuming. Frontline clinicians in the United States suggested some options for streamlining the login process, including facial recognition for login, the ability to stay logged in and/or the ability to save username and password on the login page. Both frontline clinicians in Thailand and the United States mentioned lighting and picture scale as potential barriers. Both parties suggested some type of template or guide to help with image flow, image angles, and image scale and believe this would make images more consistent. A frontline clinician also mentioned that at times he would have trouble focusing the image,

> *For some reason, when I was taking photos, I felt like I couldn't focus on what I wanted to focus on. Sometimes, I would take multiple pictures of the same thing and try to get a really good picture, but I couldn't get a clear image of things.*

End-line users had similar feelings to the frontline clinicians as far as the feasibility of the mHealth data management platform. Generally speaking, end-line users mentioned the platform itself was feasible, simple to use, user-friendly and straightforward. Regarding the mHealth platform, a dentist end-line user stated, "I think it's something that would be kind of intuitive to a new user."

A dental specialist and a general dentist mentioned that if a charting error was made, there was not a simple way to correct the mistake. They found the easiest way to correct it was to exit out of the platform and start over. They mentioned the importance of fixing this, because they believe teledentistry should be fairly efficient due to the volume one may be reviewing. A dental therapist end-line user stated that the most difficult thing was that the platform could be slow at times. Despite the hardships, overall end-line users found the learning curve to be manageable.

Perceived Usefulness

As a whole, frontline clinicians perceived the mHealth app to be useful. Several frontline clinicians believed that the mHealth app was beneficial, particularly for increasing access to dental care and reducing the burden on emergency services. Frontline clinicians based in the United States commented, “People might not have access to dental services right away following trauma.” They also stated, “I believe this app is very useful and beneficial, especially like when kids get trauma and they don’t have a dental home, it gives them quick access to care.” Frontline clinicians in Thailand also thought that the app would be a game changer for teledentistry. Specifically, they thought the mHealth app was great for those in remote areas where they do not have a dental specialist.

Frontline clinicians in both countries did suggest they think the app would be more useful if it could include more information, such as patient age, radiographs, and better image quality. They also felt some limitations with its current scope of only TDI and suggested widening it to a broader and more general dentistry scope.

End-line users deemed the mHealth data management platform useful for teledentistry, especially for initial assessments. They mentioned its potential for timely diagnosis and usefulness when there is not a dental provider in the nearby area.

Similar to the frontline clinicians, end-line users also felt having more patient information, such as pain, chief complaint, chart notes, etc, would be beneficial and useful. They also mentioned that suboptimal image quality of some dental photos may hinder the diagnosis of minor trauma and recommended that this be addressed for future use.

Compatibility

Both frontline clinicians and end-line users felt that the mHealth app and platform were compatible with the way existing dental cases and TDI cases are treated. Meaning both parties felt the technology was not too out of the realm and could be useful for handling trauma cases. A frontline clinician and an end-line user eluded to the fact that “as a society we use technology every day, and some of the younger generations rely on technology so we are already familiar with it.” They stated that most people are using smartphones every day, so there likely would not be a learning curve for widespread TDI-based teledentistry. Future improvements for the mHealth app and platform, as recommended by the interviewees, included fine-tuning of the odontogram and patient chart displayed on the platform, and integration of the mHealth model with the practice management software available in the market.

Self-Image and Social Influences

The frontline clinicians expressed that using the mHealth app did not significantly alter their self-image or how they were perceived by others. They felt the app was generally viewed positively as a helpful tool in the dental care process. A frontline clinician in Thailand stated, “there have been no bad attitudes from superiors when using the application.”

Although frontline clinicians generally had a positive perception of using an mHealth app, some concerns were raised about patients questioning the provider’s competence if they relied heavily on the app for diagnosis. One participant mentioned, “If the parent sees a provider using the application on their child they may question their intelligence as to why they are not confident in diagnosing on their own.” They also alluded that older generations may potentially have similar concerns.

Four out of 5 end-line users felt that using a teledentistry platform, such as mHealth, would have a positive impact on their professional images. They mentioned it would showcase their adaptability to using new technologies. The other end-line user, a general dentist, thought that professionals in the field would not view him or his work any differently whether he used the technology or not.

The only concern mentioned was initially, there may be some pushback, as health professionals have seen this when implementing other new technologies until the entire team is familiar and comfortable with it.

Self-Efficacy

Frontline clinicians felt confident using the mHealth app after the initial training, citing their familiarity with smartphone technology as a contributing factor. One frontline clinician stated, “Once we had the training, I was very comfortable with how to utilize the technology and then since I’ve used my phone a lot and taken a lot of images before, it was pretty streamline for me to be able to use.” Another made a similar comment,

> *I felt pretty comfortable. I felt familiar with using a lot of apps and utilizing that app wasn’t very challenging, and then taking the images were pretty easy.*

All end-line users also felt confident in their ability to use the mHealth platform effectively after receiving the initial training. End-line users attributed their confidence to the training session and prior experiences with other electronic charting and dental record systems. There were no barriers or negative experiences in regards to self-efficacy reported by end-line users.

Voluntariness and Behavioral Intention

Frontline clinicians indicated a willingness to use the app voluntarily, recognizing its potential benefits in improving access to care and streamlining the diagnostic process. One participant said, “I think it’s beneficial because, in terms of increasing access, people might not always have access to dental services right away following trauma.”

They also suggested the app could be more widely adopted if it included additional features and addressed current limitations, such as the need for frequent logins and better image quality mentioned earlier.

There was a strong willingness among end-line users to adopt the mHealth platform for future use, especially for screening and remote consultations. Most of them saw it as a great triage tool to get their patients to the right place in a timely manner. End-line users saw it as a valuable tool for patient care and were open to integrating it into their daily practices, despite some initial hesitations. Per one end-line user, “I would definitely be willing to use it, I think it’s a very valuable asset.” Another end-line user stated, “I’d say I’m very likely to use it in the future.”

One end-line user, the dental specialist, was concerned with how the app would integrate into other dental practice software in an attempt to have more patient information available for the reviewer. Another end-line user, a general dentist, thought it would be beneficial to be able to communicate with the patient through the app, whether that be through voice, FaceTime (Apple Inc), or text messaging.

Anxiety

The frontline clinicians overall seemed comfortable with using and capturing photos on the smartphones, as they use them regularly. Initial anxiety among frontline clinicians was noted due to the unfamiliarity with using the specific mobile app for dental assessments. Their anxiety lessened over time as they became more familiar with the process. Those frontline clinicians witnessed a wide range of anxiety among patients and state fairgoers who were getting their dental photos taken. The most anxiety from patients was reported in a Minnesota community clinical setting. Frontline clinicians attributed this mostly to privacy issues, stating that some of the patients are potentially illegal immigrants, and became hesitant once they found out their picture would be acquired. However, anxiety decreased among some patients once they fully understood the study purpose and applications as well as the photos taken were limited to their mouths. Minnesota state fairgoers were less anxious, as they had agreed to come to the research facilities and learn more about the study, and willingly came in to learn about the study, versus being at the dentist for an oral health issue. The frontline clinicians in Thailand did not notice much anxiety from their patients. They credited this to previous use of intraoral cameras during dental examinations and the fact that participants had previously gone through and provided an informed consent.

When it came to using the mHealth platform, end-line users were overall comfortable and felt the dental charting they did was straightforward. They did report anxiety when it came to system charting glitches, such as forgetting which teeth they had already marked as sound, and the accuracy of their diagnoses when the quality of some images was less than ideal.

Facilitating Conditions

Frontline clinicians appreciated the training sessions and written materials provided, which helped them navigate the app effectively. One frontline clinician in the United States stated,

> *I think the research team did a good job training us through this, the app, and showing us how to use it, and they even did a one-on-one calibration when we got the phones from them.*

The same interviewee thought that a couple of practice sessions and being provided with model images would have helped them have a better understanding and emulate the photo expectations. The frontline clinicians based in Thailand thought a video tutorial would have helped them better understand what angles they should take photos from. They found the most difficult part was needing to log in to the app and connect to the Wi-Fi every time they went to use the app.

When it came to resources, knowledge, or assistance, end-line users emphasized the importance of adequate training and support to maximize the platform's effectiveness. They felt the training they received from the research team was adequate, and liked the guidelines so everyone was working off the same standards. However, an end-line user suggested that going through a couple of cases together would have been even more of help.

As stated earlier in this article, end-line users felt that including more patient information, or integrating with patients' dental practice software to have access to patient dental history would make the diagnosis process easier.

Attitudes Toward a Behavior

Overall, the frontline clinicians had a positive attitude toward using the mHealth app, as it was straightforward to use and they recognized its potential to enhance dental care delivery. They thought that if this dental model were integrated more widely, it could make a significant difference, especially in areas with a shortage of dental health professionals. The mHealth model was seen as a step forward in leveraging technology to improve health care access and efficiency.

Frontline clinicians thought the experience could be even more positive by suggesting things, such as streamlining the technology by automating the photo-taking process to save and switch to the next photo without manual intervention, and user interface enhancements, such as making the app more child-friendly with colors or cartoons.

Overall, the end-line users' attitudes toward using the mHealth platform were positive. End-line users acknowledged its potential to streamline workflows and improve patient outcomes. They appreciated the convenience and accessibility it offered. Specifically, they noted the value of having standardized views for each patient and appreciated the ability to zoom in on the photos and the ability to chart multiple conditions on the tooth.

Despite the positive attitudes toward the platform and use, end-line users recognized some areas of improvement and future development. One suggestion in an attempt to standardize the diagnosis was to integrate the trauma guidelines into the platform for quick reference. Another end-line user suggested making the reviewing platform into an app that could be easily accessed with a smartphone to improve convenience during emergency situations.

Discussion

Principal Findings

This study revealed participant perspectives toward the mHealth app and platform. Participants highlighted improved and faster diagnosis, underscoring perceived usefulness, and generally found the system easy to use despite login and photo clarity challenges. The app and platform seemed compatible with current dental practices. Most users responded positively, though some worried patients might perceive overreliance on technology. Training increased confidence and reduced anxiety, leading to stronger self-efficacy. Some were ready to keep using the app, while others wanted improvements first, showing mixed intentions. Overall, users expressed positive attitudes toward its role in TDI assessment.

The 9 themes illustrate how TAM constructs operate in practice. Perceived usefulness related to faster diagnosis and improved access, while ease of use was reflected in straightforward operation. Compatibility was evident through integration with existing practices, and self-efficacy was reinforced by training and calibration. Social influences appeared in concerns about patient perceptions and peer resistance, while voluntariness and behavioral intention were expressed in willingness or reluctance to continue use. Anxiety, facilitating conditions, and attitudes toward use further shaped adoption, underscoring the interplay between technical usability and human factors.

When implementing teledentistry in clinical settings, its impacts on oral health outcomes [35], patient satisfaction [36], and cost-effectiveness [35] are commonly taken into account, and these effects have been regularly reported in literature [16,17,20,35,36]. Despite our belief in patient-centered care [37,38], health care workers' roles are central to mHealth adoption. In this study, smartphone-based photography was regarded as feasible, acceptable, and usable for remote TDI assessment. The results suggested potential applications in the initial assessment of emergency cases (such as trauma) and in triage or screening contexts (such as dental caries), warranting empirical testing in settings serving underserved populations and within interprofessional emergency workflows [2]. A prior caregiver study using the same mHealth app for dental screening and caries assessment in children also supported its feasibility [14], but direct comparisons are limited and require confirmatory research. Several other studies have also highlighted the relevance of including frontline and end-line users' perspectives in the determination of feasibility, acceptability, and usability of mHealth technologies for dental care and oral health promotion [13,15-18,20]. Those studies reported benefits including diagnostic accuracy [13], communication [17], patient-provider relationship [17], efficiency [18], and medication adherence [20]. Compellingly, some users of a non–photography-related mHealth app recommended adding a photo-capture-and-forward function to it [15], which happens to be a key characteristic of our mHealth model. This could also imply the viability of smartphone-based photography in the mHealth modality of teledentistry.

Training and calibration were key facilitators for the use of this mHealth model, increasing comfort and confidence among both frontline clinicians and end-line users. Literature suggested that adequate training decreases uncertainty, strengthens self-efficacy, and enhances mHealth usability [10,15,39,40]. A recent scoping review has found approximately 20 teledentistry educational programs [41]. Building on this evidence, curricula should incorporate topics such as virtual communication and assessment, technology troubleshooting, billing, and ethics, to help learners develop teledentistry competencies. These could be complemented by simulated training and clinical practice [42]. Because familiarity enhances usability, periodic and continuous training is recommended to refresh skills and sustain competence.

Patient and clinician perceptions may pose barriers. Some older patients might view smartphone use during diagnosis as a sign of incompetence, while younger or tech-savvy patients were more accepting [43]. Education to improve digital literacy and highlight benefits can support adoption. Resistance from colleagues may also hinder uptake. Inclusive development, professional development, enhanced communication, and reduced operational burdens can facilitate acceptance [44].

This study is exploratory and focused on user perceptions rather than on product design or feature validation; nevertheless, participants also proposed technical suggestions. Consistent image quality is essential, and camera grids or templates could help. Integration of electronic patient records would strengthen utility, while governance and ethical frameworks addressing data linkage, artificial intelligence (AI), and security are needed [45,46]. Biometric or voice login could reduce screen contact while preserving security. Finally, a risk-averse "ethical brake" approach—safeguards that trigger additional verification before data sharing or high-stakes decisions—has been proposed to inform the design of login, data sharing, and reviewer feedback loops, helping minimize potential harms as mHealth and AI features are introduced [46].

To ensure trustworthiness, investigator and data source triangulation were used to test the validity [47]. The interview questions were created and reviewed by a multimember researcher panel. Two moderators and an observer were jointly involved in data collection and analysis. These approaches broadened the focus of this research and reduced researcher biases. Additionally, the data were collected from the participants of different roles (frontline clinicians and end-line users), qualifications (dental specialists, general dentists, dental therapists, and dental students), sexes (females and males), and locations (Thailand and the United States). This participant pool contributed to diverse perspectives and validation of data.

Using a focus group interview method presents certain limitations. Group composition was based on participants' availability rather than self-selection, which may have influenced dynamics. As noted in prior literature, dominant voices can shape group discourse [48]. To mitigate this, moderators were trained to foster inclusive participation, and follow-up emails were sent to all participants to invite additional reflections. Cultural differences between Thailand and the United States also warrant consideration, even though interviews were conducted separately by country. Prior research on communication styles suggested Thais might be more inclined to accommodate audience expectations, whereas Americans might prioritize balancing personal goals with audience interests [49]. These differing cultural norms could have influenced how frontline clinicians and end-line users expressed their views. Furthermore, because interviews with Thai clinicians were conducted in English rather than their primary language, some nuance may have been lost in cross-cultural expression; however, presession proficiency checks, member checking, and transcript review by a bilingual researcher helped mitigate this risk.

While no participants were selected or excluded by demographic or professional criteria, the small sample size remains a limitation. The operation of this mHealth model required 11 frontline clinicians to take dental photos and 5 end-line users to diagnose TDI; all except 1 frontline clinician (a coauthor) participated in this qualitative study. Limited generalizability due to small samples in telemedicine and telehealth research has been noted in an integrated review [50]. Nevertheless, thematic saturation was reached early, with no new codes emerging after the fourth interview, and data collection concluded after 7 sessions. This suggests that the dataset adequately captured the range of user perspectives for this workflow, consistent with Guest et al [51] findings that as few as 6 interviews may be sufficient for thematic analysis and interpretation. Future work should include larger and more culturally diverse clinician samples and explore similar mHealth models to enhance transferability.

The incentive differences between countries stemmed from institutional funding restrictions, which prohibit disbursing school-supported funds to individuals not legally eligible to work in the United States. Although they did not affect participation rates, such disparities may have introduced subtle response biases or influenced perceived power dynamics. To promote equitable participation in future international research, investigators should consider globally inclusive funding mechanisms or secure domestic support within each participating country.

Interpersonal dynamics [52] and prior researcher-participant familiarity [53] may also influence participant responses. To mitigate potential courtesy bias and power dynamics arising from interviewer familiarity, reflexivity and bracketing strategies were used throughout the study. Moderators received training to foster inclusive dialogue, actively invite dissenting views, and minimize hierarchical influence during interviews. A structured interview guide (Table 1) was used to ensure consistency across sessions and reduce the influence of moderator bias or familiarity on discussion flow. Although the PI observed all sessions, he did not moderate or participate; his webcam and microphone remained off throughout. Participants were not supervised, assessed, or taught by the PI during or after the study period. Their feedback, whether positive or critical, was valued equally and anonymized during reporting to encourage candor. Follow-up emails invited dissenting views, further reducing social pressure and courtesy bias. Prior researcher-participant relationships have also been identified as a potential asset in qualitative research, helping to build rapport and facilitate open dialogue [53].

Future research should incorporate role-specific interview elements, mixed methods approaches, and patient perspectives, alongside cost-effectiveness analyses. With evidence supporting remote assessment of TDI [10], dental caries [11,14], and oral cancer [12], diagnostic capacities could expand to other oral conditions such as tooth wear [54], developmental defects [55], malocclusion [56], and temporomandibular disorders [57].

Scaling mHealth tools in emergency dental settings requires technical and training enhancements as well as policy and infrastructure support. Regulatory frameworks for teledentistry differ across jurisdictions and may lack explicit guidance for photographic diagnosis, raising concerns around licensure, liability, and documentation standards. Reimbursement for asynchronous services—such as image-based consultations—remains inconsistent, potentially limiting uptake in emergency departments and underserved communities. Moreover, integration with existing electronic health record systems and secure data transmission protocols is vital to ensure interoperability, privacy compliance, and operational efficiency. Addressing these regulatory, reimbursement, and infrastructure barriers is essential to move beyond pilot models and establish sustainable, scalable mHealth solutions.

Equally important is the role of educational and professional development in supporting adoption and innovation. Integrating photography-based mHealth workflow into emergency care requires structured training and clear protocols for photo capture and clinical assessment. Competency development in teledentistry should be incorporated into continuing education and dental curricula to prepare both current and future providers. At the system level, health care organizations can leverage the insights from this study to guide infrastructure planning, resource allocation, and policy for effective integration of mHealth technologies.

Conclusion

Within the study limitations, frontline clinicians and end-line users regarded smartphone-based photography as a feasible, acceptable, and usable mHealth tool for remote dental assessment. Strengthening self-efficacy through ongoing education and professional development will be essential, while patient and clinician resistance may be mitigated by strategies such as inclusive development, communication, reduction of operational burdens, and targeted training. Technical refinements, including standardized photo templates, integration with patient management systems, and streamlined login, could further enhance usability. Future research should pilot this model in larger and more diverse clinical populations, evaluate cost-effectiveness, and explore its integration into broader oral health workflows beyond TDI.

Acknowledgments

The authors acknowledge Minnesota Masonic Charities, International Association for Dental Research, the University of Minnesota Center for Global Health and Social Responsibility, and the University of Minnesota School of Dentistry for funding support. We also thank the research participants and university faculty and staff who made this possible through their involvement. Part of this work has been presented in MS’s Honors thesis [58]. A publicly available preprint of this work was posted online following the manuscript’s submission to the journal for publication [59].

Funding

This study was supported with a Department of Pediatrics Driven to Discover Grant by Minnesota Masonic Charities, a Regional Development Program Grant by the International Association for Dental Research, a Global Health Seed Grant by the University of Minnesota Center for Global Health and Social Responsibility, and a Faculty Start-up Funding for BH by the University of Minnesota School of Dentistry.

Data Availability

The datasets generated or analyzed during this study are not publicly available because participants of this study did not give written consent for their recordings to be publicly shared, but the transcribed data are available from the corresponding author (BH) on reasonable request.

Conflicts of Interest

None declared.

References

1. Huang B, Marcenes W, Croucher R, Hector M. Activities related to the occurrence of traumatic dental injuries in 15- to 18-year-olds. Dent Traumatol 2009;25(1):64-68. [doi: 10.1111/j.1600-9657.2008.00685.x] [Medline: 19208012]
2. Hall E, Hickey P, Nguyen-Tran T, Louie J. Dental Trauma in a pediatric emergency department referral center. Pediatr Emerg Care 2016;32(12):823-826. [doi: 10.1097/PEC.0000000000000961] [Medline: 27898626]
3. Schiappa M, Palla B, Callahan NF. What are the incidence and costs associated with interfacility emergency department transfer patients treated by oral and maxillofacial surgeons? Oral Surg Oral Med Oral Pathol Oral Radiol 2024;138(5):608-612. [doi: 10.1016/j.oooo.2024.06.008] [Medline: 39217075]
4. Eggmann F, Haschemi AA, Doukoudis D, Filippi A, Verna C, Walter C, et al. Impact of the COVID-19 pandemic on urgent dental care delivery in a Swiss university center for dental medicine. Clin Oral Investig 2021;25(10):5711-5721 [FREE Full text] [doi: 10.1007/s00784-021-03872-1] [Medline: 33710460]
5. Namakian M, Subar P, Glassman P, Quade R, Harrington M. In-person versus "virtual" dental examination: congruence between decision-making modalities. J Calif Dent Assoc 2012;40(7):587-595. [Medline: 22916380]
6. Davies A, Mueller J. Developing Medical Apps and mHealth Interventions. Cham, Switzerlad: Springer; 2020.
7. Harris B, Ajisola M, Alam RM, Watkins JA, Arvanitis TN, Bakibinga P, et al. Mobile consulting as an option for delivering healthcare services in low-resource settings in low- and middle-income countries: a mixed-methods study. Digit Health 2021;7:20552076211033425 [FREE Full text] [doi: 10.1177/20552076211033425] [Medline: 34777849]
8. Mariño RJ, Uribe SE, Chen R, Schwendicke F, Giraudeau N, Scheerman JFM. Terminology of e-oral health: consensus report of the IADR's e-Oral Health Network Terminology Task Force. BMC Oral Health 2024;24(1):280 [FREE Full text] [doi: 10.1186/s12903-024-03929-z] [Medline: 38419003]
9. Xiao J, Meyerowitz C, Ragusa P, Funkhouser K, Lischka TR, Mendez Chagoya LA, National Dental Practice-Based Research Network Collaborative Group. Assessment of an innovative mobile dentistry ehygiene model amid the covid-19 pandemic in the national dental practice-based research network: protocol for design, implementation, and usability testing. JMIR Res Protoc 2021;10(10):e32345 [FREE Full text] [doi: 10.2196/32345] [Medline: 34597259]
10. Huang B, Estai M, Pungchanchaikul P, Quick K, Ranjitkar S, Fashingbauer E, et al. Mobile health assessment of traumatic dental injuries using smartphone-acquired photographs: a multicenter diagnostic accuracy study. Telemed J E Health 2024;30(10):2592-2600 [FREE Full text] [doi: 10.1089/tmj.2024.0012] [Medline: 38946689]
11. Estai M, Kanagasingam Y, Huang B, Checker H, Steele L, Kruger E, et al. The efficacy of remote screening for dental caries by mid-level dental providers using a mobile teledentistry model. Community Dent Oral Epidemiol 2016;44(5):435-441. [doi: 10.1111/cdoe.12232] [Medline: 27111291]
12. Birur PN, Sunny SP, Jena S, Kandasarma U, Raghavan S, Ramaswamy B, et al. Mobile health application for remote oral cancer surveillance. J Am Dent Assoc 2015;146(12):886-894. [doi: 10.1016/j.adaj.2015.05.020] [Medline: 26610833]
13. Abuabara A, de Castro J, Locks M, Pezzin A, Mattos N, de Araújo CM, et al. Evaluation of Endo 10 mobile application as diagnostic tool in endodontics. J Clin Exp Dent 2023;15(8):e612-e620 [FREE Full text] [doi: 10.4317/jced.60342] [Medline: 37674608]
14. Azimi S, Fernando C, Estai M, Patel J, Silva D, Tennant M. Experience of primary caregivers in utilising an mHealth application for remote dental screening in preschool children. Aust Health Rev 2023;47(5):545-552. [doi: 10.1071/AH23110] [Medline: 37580061]
15. Ibarra-Noriega A, Yansane A, Mullins J, Simmons K, Skourtes N, Holmes D, et al. Evaluating and improving the usability of a mHealth platform to assess postoperative dental pain. JAMIA Open 2024;7(1):ooae018 [FREE Full text] [doi: 10.1093/jamiaopen/ooae018] [Medline: 38476372]
16. Van Cleave JH, Fu MR, Bennett AV, Concert C, Riccobene A, Tran A, et al. The usefulness of the Electronic Patient Visit Assessment (ePVA) as a clinical support tool for real-time interventions in head and neck cancer. Mhealth 2021;7:7 [FREE Full text] [doi: 10.21037/mhealth-19-250] [Medline: 33634190]

17. Tokede B, Yansane A, Ibarra-Noriega A, Mullins J, Simmons K, Skourtes N, et al. Evaluating the impact of an mHealth platform for managing acute postoperative dental pain: randomized controlled trial. JMIR Mhealth Uhealth 2023;11:e49677 [FREE Full text] [doi: 10.2196/49677] [Medline: 37933185]
18. Kalenderian E, Tungare S, Mehta U, Hamid S, Mungia R, Yansane A, et al. Patient and dentist perspectives on collecting patient reported outcomes after painful dental procedures in the National Dental PBRN. BMC Oral Health 2024;24(1):201 [FREE Full text] [doi: 10.1186/s12903-024-03931-5] [Medline: 38326805]
19. Del Carmen MDC, Cagigas-Muñiz D, García-Robles R, Oprescu AM. Reducing dental anxiety in children using a mobile health app: Usability and user experience study. JMIR Form Res 2023;7:e30443 [FREE Full text] [doi: 10.2196/30443] [Medline: 37889521]
20. Pereira-Salgado A, Westwood JA, Russell L, Ugalde A, Ortlepp B, Seymour JF, et al. Mobile health intervention to increase oral cancer therapy adherence in patients with chronic myeloid leukemia (The REMIND System): clinical feasibility and acceptability assessment. JMIR Mhealth Uhealth 2017;5(12):e184 [FREE Full text] [doi: 10.2196/mhealth.8349] [Medline: 29212628]
21. Almazrooa SA, Mansour GA, Alhamed SA, Ali SA, Akeel SK, Alhindi NA, et al. The application of teledentistry for Saudi patients' care: a national survey study. J Dent Sci 2021;16(1):280-286 [FREE Full text] [doi: 10.1016/j.jds.2020.04.014] [Medline: 33384810]
22. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-339.
23. Wallis L, Blessing P, Dalwai M, Shin SD. Integrating mHealth at point of care in low- and middle-income settings: the system perspective. Glob Health Action 2017;10(sup3):1327686 [FREE Full text] [doi: 10.1080/16549716.2017.1327686] [Medline: 28838302]
24. Scauso M. Interpretivism: Definitions, Trends, and Emerging Paths. Oxford, UK: Oxford Research Encyclopedia of International Studies; 2024.
25. 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]
26. Health Insurance Portability and Accountability Act of 1996. Office of the Assistant Secretary for Planning and Evaluation. 1996. URL: https://aspe.hhs.gov/reports/health-insurance-portability-accountability-act-1996 [accessed 2025-11-21]
27. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]
28. Klingberg A, Sawe HR, Hammar U, Wallis LA, Hasselberg M. m-Health for burn injury consultations in a low-resource setting: an acceptability study among health care providers. Telemed J E Health 2020;26(4):395-405 [FREE Full text] [doi: 10.1089/tmj.2019.0048] [Medline: 31161967]
29. Moore GC, Benbasat I. Development of an instrument to measure the perceptions of adopting an information technology innovation. Inf Syst Res 1991;2(3):192-222. [doi: 10.1287/isre.2.3.192]
30. Compeau D, Higgins C. Computer self-efficacy: development of a measure and initial test. MIS Q 1995;19(2):189-211.
31. Fishbein M, Ajzen I. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Boston, USA: Addison-Wesley; 1975.
32. Thompson RL, Higgins CA, Howell JM. Personal computing: toward a conceptual model of utilization. MIS Q 1991;15(1):125-143.
33. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. Int J Qual Methods 2023;22(11). [doi: 10.1177/16094069231205789]
34. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Scheerman JFM, Qari AH, Varenne B, Bijwaard H, Swinckels L, Giraudeau N, et al. A systematic umbrella review of the effects of teledentistry on costs and oral-health outcomes. Int J Environ Res Public Health 2024;21(4):407 [FREE Full text] [doi: 10.3390/ijerph21040407] [Medline: 38673320]
36. Emami E, Harnagea H, Shrivastava R, Ahmadi M, Giraudeau N. Patient satisfaction with e-oral health care in rural and remote settings: a systematic review. Syst Rev 2022;11(1):234 [FREE Full text] [doi: 10.1186/s13643-022-02103-2] [Medline: 36309732]
37. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. Med Care Res Rev 2013;70(4):351-379. [doi: 10.1177/1077558712465774] [Medline: 23169897]
38. Howell SE, Fukuoka B. Teledentistry for patient-centered screening and assessment. Dent Clin North Am 2022;66(2):195-208. [doi: 10.1016/j.cden.2022.01.002] [Medline: 35365273]
39. National Research Council. In: Druckman D, Bjork RA, editors. Learning, Remembering, Believing: Enhancing Human Performance. USA: National Academies Press; 1994.
40. Birur N, Patrick S, Bajaj S, Raghavan S, Suresh A, Sunny S, et al. A novel mobile health approach to early diagnosis of oral cancer. J Contemp Dent Pract 2018;19(9):1122-1128 [FREE Full text] [Medline: 30287715]
41. Bakhshaei A, Ramachandran S, Brondani M. Teledentistry within oral health care providers' training: a scoping review. Eur J Dent Educ 2024;28(2):631-644. [doi: 10.1111/eje.12990] [Medline: 38279780]

42. McLeod CD, Adatorwovor R, Brame JL, White BA, Weintraub JA. Teledentistry: dental hygiene students' knowledge, attitudes, and curriculum recommendations. J Dent Hyg 2021;95(5):64-72. [Medline: 34654717]
43. Giles-Smith L, Spencer A, Shaw C, Porter C, Lobchuk M. A study of the impact of an educational intervention on nurse attitudes and behaviours toward mobile device use in hospital settings. J Can Health Libr Assoc 2017;38(1):12-29.
44. Safi S, Thiessen T, Schmailzl KJ. Acceptance and resistance of new digital technologies in medicine: Qualitative study. JMIR Res Protoc 2018;7(12):e11072 [FREE Full text] [doi: 10.2196/11072] [Medline: 30514693]
45. 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 2022:1-14 [FREE Full text] [doi: 10.1007/s41347-022-00242-w] [Medline: 35194561]
46. Thurzo A, Thurzo V. Embedding fear in medical AI: A risk-averse framework for safety and ethics. AI 2025;6(5):101. [doi: 10.3390/ai6050101]
47. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014;41(5):545-457. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
48. Leung FH, Savithiri R. Spotlight on focus groups. Can Fam Physician 2009;55(2):218-219 [FREE Full text] [Medline: 19221085]
49. Knutson TJ, Komolsevin R, Chatiketu P, Smith VR. A cross-cultural comparison of Thai and US American rhetorical sensitivity: implications for intercultural communication effectiveness. Int J Intercult Relat 2003;27(1):63-78.
50. Groom LL, McCarthy MM, Stimpfel AW, Brody AA. Telemedicine and telehealth in nursing homes: an integrative review. J Am Med Dir Assoc 2021;22(9):1784-1801.e7 [FREE Full text] [doi: 10.1016/j.jamda.2021.02.037] [Medline: 33819450]
51. Guest G, Bunce A, Johnson L. How many interviews are enough? Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822x05279903]
52. de Souza J, Gillett K, Salifu Y, Walshe C. Changes in participant interactions. Using focus group analysis methodology to explore the impact on participant interactions of face-to-face versus online video data collection methods. Int J Qual Methods 2024;23(2). [doi: 10.1177/16094069241241151]
53. Roiha A, Iikkanen P. The salience of a prior relationship between researcher and participants: reflecting on acquaintance interviews. Res Method Appl Linguistics 2022;1(1):100003. [doi: 10.1016/j.rmal.2021.100003]
54. Lee J, Fung S, Yong R, Ranjitkar S, Kaidonis J, Evans AR, et al. Tooth wear development in the Australian Aboriginal dentition from Yuendumu: a longitudinal study. PLoS One 2021;16(7):e0254151 [FREE Full text] [doi: 10.1371/journal.pone.0254151] [Medline: 34242291]
55. Pitiphat W, Luangchaichaweng S, Pungchanchaikul P, Angwaravong O, Chansamak N. Factors associated with molar incisor hypomineralization in Thai children. Eur J Oral Sci 2014;122(4):265-270. [doi: 10.1111/eos.12136] [Medline: 24924351]
56. Tanny L, Huang B, Shaweesh A, Currie G. Characterisation of anterior open bite in primary school-aged children: a preliminary study with artificial neural network analysis. Int J Paediatr Dent 2021;31(5):576-582. [doi: 10.1111/ipd.12759] [Medline: 33222319]
57. Huang B, Takahashi K, Sakata T, Kiso H, Sugai M, Fujimura K, et al. Increased risk of temporomandibular joint closed lock: a case-control study of ANKH polymorphisms. PLoS One 2011;6(10):e25503 [FREE Full text] [doi: 10.1371/journal.pone.0025503] [Medline: 22003394]
58. Shenouda M. Dental photographer perception of participant attitude with the use of teledentistry [Thesis]. Minneapolis, MN, USA: University of Minnesota; 2024.
59. Schultz EC, Huang B, Shenouda M, Estai M, Ranjitkar S, Louie JP, et al. In-Review: Perspectives of Front-Line Clinicians and Remote Reviewers on Smartphone-Based Photography for Assessing Traumatic Dental Injuries: A Qualitative Study. United States: University of Minnesota URL: https://conservancy.umn.edu/items/51c73533-eb65-4257-a6c2-33895147b00b [accessed 2026-01-06]

Abbreviations

AI: artificial intelligence
COREQ: Consolidated Criteria for Reporting Qualitative Research
CSIRO: Commonwealth Scientific and Industrial Research Organisation
HIPAA: Health Insurance Portability and Accountability Act
mHealth: mobile health
PI: principal investigator
TAM: Technology Acceptance Model
TDI: traumatic dental injury

Edited by A Kushniruk; submitted 19.Aug.2025; peer-reviewed by R Musa, A Thurzo; comments to author 12.Sep.2025; revised version received 15.Dec.2025; accepted 02.Jan.2026; published 26.Jan.2026.

Please cite as:
Schultz EC, Huang B, Shenouda M, Estai M, Ranjitkar S, Louie JP, Pungchanchaikul P
Perspectives of Frontline Clinicians and End-Line Users on Smartphone-Based Photography for Assessing Traumatic Dental Injuries: Focus Group Interview Study and Thematic Analysis
JMIR Hum Factors 2026;13:e82668
URL: https://humanfactors.jmir.org/2026/1/e82668
doi:10.2196/82668
PMID:

Original Paper

Challenges and Benefits of Virtual Reality in Home-Based Poststroke Rehabilitation: Co-Design Qualitative Study

Trust Saidi[1], PhD; Ann Marie Hestetun-Mandrup[2,3], MSc; Nenad Pavel[4], PhD; Ingvild Kristina Hurum Rosseland[5], MBA; Kathinka Granum Selmer-Olsen[5], BSc; Nora Synnøve Mørk[5], MHS; Åse Bergheim[5], MSc; Minna Annika Pikkarainen[1,2], PhD

[1]Department of Product Design, Faculty of Technology, Art and Design, OsloMet – Oslo Metropolitan University, Oslo, Norway
[2]Department of Rehabilitation and Health Technology, Faculty of Health Sciences, OsloMet – Oslo Metropolitan University, Oslo, Norway
[3]Sunnaas Rehabilitation Hospital, Norway, Bjørnemyr, Norway
[4]Department of Product Design, Faculty of Technology, Art and Design, OsloMet – Oslo Metropolitan University, OSLO, Norway
[5]Forsterket rehabilitering Aker, Helseetaten Oslo Kommune, Oslo, Norway, Oslo Municipality, Aker Hospital, Oslo, Norway

Corresponding Author:
Trust Saidi, PhD
Department of Product Design
Faculty of Technology, Art and Design
OsloMet – Oslo Metropolitan University
PO Box 4
St. Olavs plass
Oslo, 0130
Norway
Phone: 47 48507071
Email: trust.saidi@oslomet.no

Abstract

Background: Stroke often leads to persistent impairments that limit daily functioning and psychosocial well-being. Virtual reality (VR) has emerged as a promising adjunct in stroke rehabilitation, although research has largely focused on clinical outcomes, with limited attention to user involvement and the experiences of multiple stakeholders in the design process.

Objective: This study aimed to explore the challenges and benefits of co-designing and using VR to support home-based poststroke rehabilitation.

Methods: A qualitative co-design case study was applied to gain an in-depth understanding of the challenges and benefits. Rapid co-design principles were used in developing VR prototypes delivered via head-mounted displays through 3 workshops with participants with stroke, health care professionals, and VR developers from November 2023 to May 2024. A design prototype revision was done based on feedback from the workshops. Data were collected via audio-taped co-design workshops with all participants and 10 successive semistructured interviews with health care professionals and VR developers conducted in a rehabilitation hospital. A thematic analysis was performed on transcribed recordings.

Results: In total, five main themes emerged: (1) adaptability for stroke-related impairments in home rehabilitation, (2) safety and ease of use, (3) goal orientation, (4) motivation, and (5) VR as a complementary tool. One of the primary challenges identified lies in the adaptability of VR systems for individuals with hemiparesis. Additionally, customization and safety concerns remain a complex barrier, as VR solutions must be capable of addressing a wide range of stroke-related impairments and aligning specific rehabilitation goals. VR demonstrated potential to enhance rehabilitation by simulating real-life tasks that encourage goal-oriented and motivating therapy. As a complementary tool, VR can enhance traditional rehabilitation by increasing the intensity and volume of therapy.

Conclusions: This study offers insight into how VR can be effectively integrated into rehabilitation practices. Its integration into rehabilitation requires alignment with established therapeutic principles within VR applications, such as adjustable task-specific training and meaningful outcomes tailored to individual needs, to ensure clinical relevance and user engagement. VR should complement, rather than replace, conventional therapy by increasing training intensity, reducing therapist workload, and extending rehabilitation into the home. Thoughtful co-design with stakeholders is key to creating VR tools that bridge the gap between structured clinical care and independent recovery, offering continuous support throughout the rehabilitation process.

(JMIR Hum Factors 2026;13:e78798) doi:10.2196/78798

KEYWORDS

virtual reality; rehabilitation; people after stroke; co-design; stroke

Introduction

Stroke ranks as the second leading cause of death globally and is a major contributor to adults' neurological and neuropsychological persistent impairments [1], impacting 17 million individuals annually [2]. This is a cause for concern as the demand for poststroke rehabilitation is expected to rise significantly [3], imposing considerable strains on both the quality of life of people after stroke and health care resources. Rehabilitation for people after stroke often requires a long period of rehabilitation and motor relearning [4]. Artificial intelligence (AI)–driven virtual reality (VR) solutions are expected to support individuals' self-managed rehabilitation and have the potential to effectively engage individuals in intensive, repetitive, and task-oriented activities [5]. They can serve as useful tools for monitoring progress and enhancing decision-making in the rehabilitation process for people after stroke and health care professionals [6].

Stroke can affect individuals to varying degrees, with motor impairments being the most frequent disability. These impairments can limit motor mobility and have a negative impact on individuals' physical activity levels and psychosocial well-being [7]. A European multicenter cohort study demonstrated that after 5 years, people after stroke experienced a decline in functional and motor outcomes, reverting to the levels observed at 2 months post stroke, highlighting that many individuals experience residual impairments [8]. Motor disabilities can vary from slight weakness to severe paralysis, impacting one's ability to perform everyday activities like eating, cooking, and dressing independently. It is estimated that about 75% of people after stroke initially face arm impairment, with only about 50% regaining arm function within 6 months after their stroke [9,10]. Although there are several studies on the effectiveness and use of VR, with some targeting specifically upper limb rehabilitation [11-13], they are mostly systematic reviews. The empirical studies on this topic are mainly based on the evaluation of the effects of VR in the rehabilitation process and not individual experiences [14-18]. While both the systematic reviews and evaluations on the effect of VR have enhanced our understanding of the benefits of these interventions, there is a paucity of knowledge on how these positive outcomes are achieved.

Experiencing a stroke often affects more than just physical activities and functions; it also impacts cognition and personal factors. The impact on life after stroke often leaves individuals feeling vulnerable and anxious about what lies ahead [19,20]. Cognitive challenges, including difficulties with memory, attention, and problem-solving, are common, as are emotional effects, such as mood swings, depression, and anxiety [21]. VR technologies targeting cognitive function and stress management have also been found to be effective [22]; however, a recognized gap in the literature exists regarding the effect that VR has on psychological well-being, which potentially affects cognitive rehabilitation. Therefore, greater attention is needed regarding how individuals recovering from stroke engage with VR solutions. This paper seeks to contribute to the growing body of knowledge in this area. Stroke rehabilitation is often a prolonged and challenging process, requiring long-term therapeutic intervention. A study by Johansen et al [23] examining the use of VR equipment in home settings for individuals with brain injuries highlighted the ongoing need for cognitive rehabilitation following hospital discharge. Participants in the study emphasized the importance of individualized VR interventions and noted that initial training typically occurred within hospital environments. These findings suggest that exploring the early implementation of VR in controlled clinical settings, facilitated through close collaboration with health care professionals, may offer valuable insights into the feasibility and effectiveness of VR-based rehabilitation post stroke.

The gaps identified above underscore the importance of understanding how VR technology can be both designed for and experienced by users within a safe and supportive environment, one that allows for a comprehensive exploration of its potential benefits as well as its inherent challenges. This demands a focus on the process of engaging the users rather than merely evaluating the effects. Focusing on the effects without elaborating on the development process poses the danger of oversimplifying the mechanisms through which VR contributes to recovery. Without a clear understanding of the underlying processes, it becomes challenging to optimize and innovate VR interventions to maximize their rehabilitation potential. Given the limited research on the use of immersive VR for stroke rehabilitation involving clinicians, people after stroke, and VR developers, it is crucial to integrate a multidisciplinary co-design approach to ensure that clinical VR solutions are both effective and user-centered [24-26]. By involving individuals with stroke, developers, and multidisciplinary health care professionals early in the design process, we can balance desired features with evidence-based design recommendations, enhancing the relevance and usability of these digital technologies. This approach also helps prevent the development of solutions that fail to meet actual treatment needs or pose safety risks to people after stroke when using VR solutions independently at home.

Even though VR systems promise more engaging experiences, especially when it comes to repetitive movements in poststroke rehabilitation, it is still not a go-to practice for many people after stroke and therapists. This study, therefore, aims to address the identified research gaps by exploring the challenges and benefits of co-designing and using VR to support home-based poststroke rehabilitation.

Methods

Overview

A qualitative co-design case study was conducted to gain an in-depth understanding of the real-life phenomenon within its environmental context [27]. A single case study approach was selected as it provided a robust means to explore "how" and "why" questions within a specific context, which gave the opportunity to investigate deeper causes of the phenomenon [28]. The data were collected through different sources of semistructured interviews and workshops.

Design Process Using Rapid Prototyping

We applied design principles and rapid prototyping [29] to develop VR prototypes (Figure 1) through 3 workshops with a mixture of participants held between November 2023 to May 2024 with the goal of (1) understanding the challenges and map the needs of people after stroke using an existing VR café scenario; (2) discuss possibilities and evaluate features in kitchen and a painting scenarios; and (3) test, evaluate, and refine the modified prototypes of the kitchen and painting scenarios through discussions and user feedback. The prototypes were immersive VR environments delivered via head-mounted displays (Oculus Quest 3 [Reality Labs]), featuring AI-driven interactive virtual characters, while other scenarios used scripted or noninteractive virtual elements with hand tracking.

To explore the potential of immersive environments in stroke rehabilitation, 3 VR scenarios, that is, a painting studio, a café, and a kitchen (Figure 2), were designed to simulate familiar and meaningful everyday settings. These environments were purposefully developed to promote creative engagement and functional interaction, allowing participants to practice activities reflecting real-life contexts in a safe and controlled virtual space. The 3 scenarios are illustrated in Figure 1.

Figure 1. Virtual reality scenarios presented to people after stroke (n=4) during stakeholder workshops. (A) Scenario 1: Painting, (B) Scenario 2: Café, and (C) Scenario 3: Kitchen.

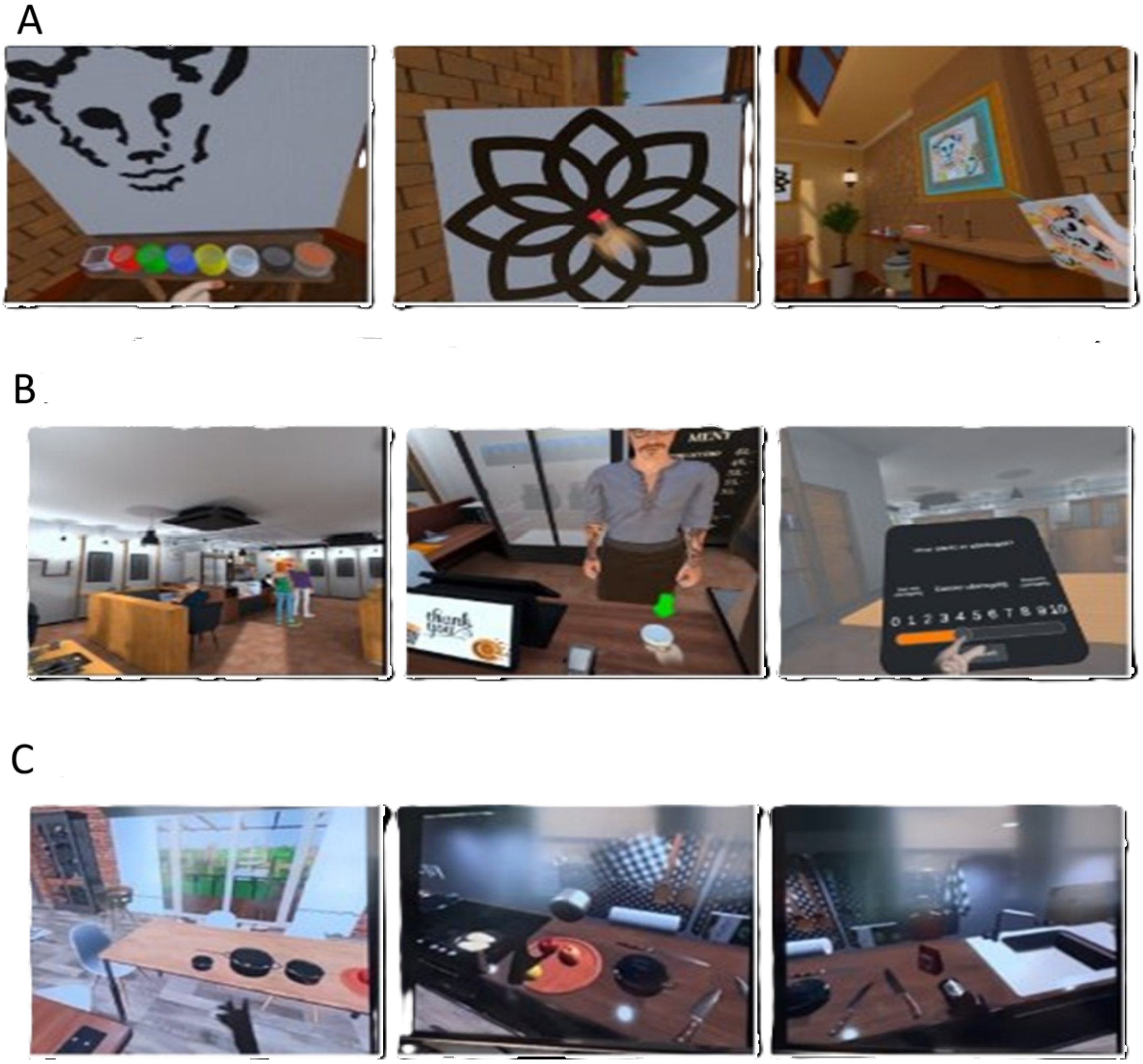
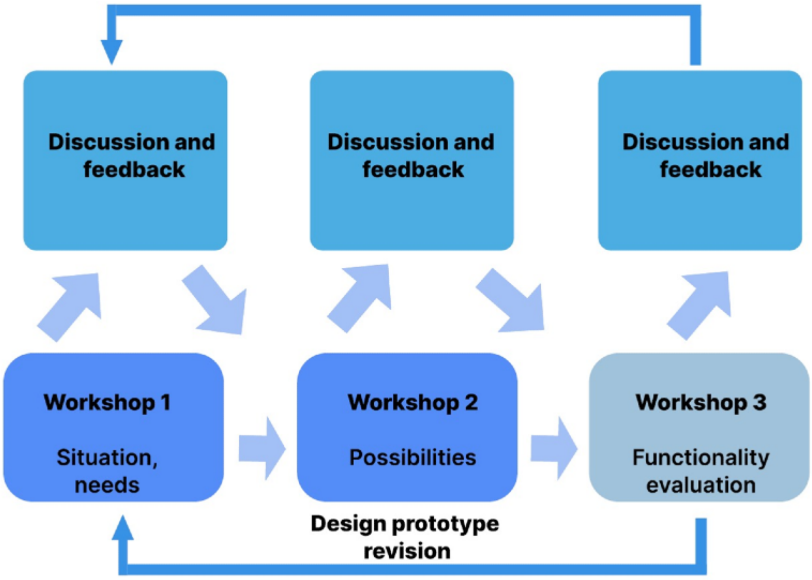


Figure 2. Flowchart illustrating the design workshops involving people after stroke, health care professionals, virtual reality developers, and researchers (n=16) to iteratively develop and refine a virtual reality prototype for poststroke rehabilitation.



The first scenario illustrates the painting environment, where the participants entered a living room and moved toward the drawing board, where equipment and colors were located. The participants could choose between freehand painting and tracing a preset illustration or figure. Guidelines for each activity were provided in all scenarios. The second scenario placed participants in a large café environment. They moved toward the cashier's desk, ordered their food, and paid using a tablet. Participants also had the opportunity to interact with other VR avatars in the café. The third scenario illustrates a kitchen environment, where the participants could move between the kitchen counter and the dining table. They could do different activities like setting the table, placing kitchen utensils in drawers, or putting apples in a bucket.

In the first workshop, painting and café scenarios were presented to people after stroke and health care professionals. For the subsequent workshops, only painting and kitchen scenarios were further developed, as these were identified as most relevant and useful for people after stroke undergoing home rehabilitation, particularly for enhancing physical and cognitive functions. The co-design process was led by 2 researchers (MAP and NP)—a professor and an associate professor—who brought insider knowledge about the co-designing and feature-adjusting phenomena. Furthermore, 2 additional researchers (TS and AMHM), a postdoctoral and doctoral candidate, provided an outsider view to challenge assumptions. This single case approach allowed researchers to follow the co-design development process and adaptation of VR systems to the needs and demands of people after stroke, health care professionals, and VR developers.

Each workshop began with a short introduction to the aim of the VR scenarios, followed by people after stroke testing the solutions, and finally followed by 2 discussion sessions with all stakeholders and 1 exclusively with health care professionals. Feedback from the workshops informed iterative revisions of the prototype, as illustrated in Figure 2.

Recruitment of Participants

The study was conducted in a rehabilitation hospital where participants were recruited to participate in the workshops from three distinct groups where the number of unique participants was: (1) health care professionals (n=10), (2) people after stroke (n=4), and (3) VR developers (n=2). All participants received previous written information and were invited to participate in both the workshops and follow-up interviews via phone or email. Health care professionals were recruited by the unit leader, while people after stroke were selected based on their involvement in ongoing rehabilitation programs within the hospital. The VR developers were already involved in organizing the workshops as part of their collaboration with the research team. They agreed to participate in interviews after the workshops were concluded. During the workshops, all 3 groups were asked if they were willing to participate in follow-up interviews. A total of 10 participants agreed, including health care professionals and VR developers. Additionally, 3 out of 4 participants with stroke agreed to engage in short conversations and provided feedback during VR testing. The 4 participants were recruited from a community-funded rehabilitation hospital in Norway based on the following inclusion and exclusion criteria: adults with a confirmed stroke diagnosis, speaking either English or Norwegian, and living at home in the subacute or chronic phase (more than 3 months since the last stroke) as defined by a stroke trajectory framework [30]. In addition, the participants needed to have a score of 18 or above on the Montreal Cognitive Assessment (MoCA) [31] and the ability to complete specific tasks from the Action Research Arm Test [32,33], such as pouring water from one glass to another and placing a hand on top of the head. Exclusion criteria included individuals younger than 18 years, pregnant women, and people with receptive or expressive aphasia or memory and communication impairments that hindered them from using the prototypes. The 4 participants with stroke had a balanced gender distribution and ranged in age from 53 to 64 years. The participants experienced different types of strokes; 2 experienced ischemic strokes, while the other 2 experienced hemorrhages (1 subarachnoid and 1 brainstem). None of the participants with stroke had previous experience with VR. Additionally, several health care professionals with expertise in stroke rehabilitation (ie, physiotherapists, occupational therapists, nurses, and medical doctors) and VR developers participated. The stakeholders, consisting of people after stroke, health care professionals, and VR developers, were involved in all 3 workshops as illustrated in Table 1.

Table 1. Overview of the participants excluding the researchers involved in the workshops (n=16) and interviews (n=10).

Stakeholder	Workshop attendance	Individual interviews
Researchers	Workshop 1: n=3 Workshop 2: n=3 Workshop 3: n=4 Total number of unique participants: n=4	—[a]
People after stroke	Workshop 1: n=3 Workshop 2: n=3 Workshop 3: n=3 Total number of unique participants: n=4	—
Health care professionals	Workshop 1: n=8 Workshop 2: n=5 Workshop 3: n=10 Total number of unique participants: n=10	Number of interviewees=8 (4 occupational therapists, 2 physiotherapists, 1 medical doctor, and 1 nurse)
VR[b] technology developers	Workshop 1: n=2 Workshop 2: n=2 Workshop 3: n=2 Total number of unique participants: n=2	Number of interviewees=2 (VR developers)

[a] Not applicable.

[b] VR: virtual reality.

Data Collection

Audio recordings from each of the 3- to 4-hour co-design workshops [34] and collaboratively developed individual interviews conducted at a rehabilitation hospital served as the primary data sources for this case study. A total of 10 semistructured interviews, lasting 60-90 minutes each, were conducted and audio-recorded. Open-ended questions were used in both workshops and interviews (refer to interview guide in Multimedia Appendix 1). All meetings were conducted face-to-face. Workshop 3, being both comprehensive and advanced, built upon the outcomes of Workshops 1 and 2 and served as the primary data source for this study. Both workshops and interviews were transcribed and anonymized, with no real names or personal information retained from participants, to ensure confidentiality and adhere to ethical research standards. Researchers from Oslo Metropolitan University, led by designers with experience in design-driven innovation, user experience design, and digital prototyping (MAP and NP), mediated the workshops, while clinicians provided support to participants with stroke during the testing. Interviews and initial analysis were conducted by 3 experienced qualitative researchers (AMHM, TS, and NP) from Oslo Metropolitan University, with backgrounds in physiotherapy, science and technology studies, and computer and design science. Tasks were assigned based on relevant expertise; for instance, Norwegian-language interviews were conducted by fluent researchers, while interviews with VR developers were handled by a researcher specializing in the science, technology, and society interface.

The decision to conclude data collection after 10 interviews was guided by the concept of information power [35], which asserts that sample adequacy in qualitative research depends on the relevance and richness of the data in relation to the study's aim. Given the focused objective of gathering actionable feedback for iterative refinement of VR prototypes rather than achieving full thematic saturation, a small, purposefully selected sample was appropriate. This approach aligns with rapid prototyping methodology, which prioritizes targeted insights to inform design decisions over exhaustive theme development [36]. The study also followed participatory research principles, prioritizing a representative sample of end users to ensure relevance and applicability. The sample size (n=10) also aligns with Creswell's [37] recommended range of 5-25 participants for phenomenological studies. Finally, triangulation [38] using audio recordings from stakeholder workshops and individual interviews enhanced the credibility and depth of the findings.

Data Analysis

The data were transcribed verbatim using the transcription software Whisper (OpenAI) and manually checked for accuracy. Thematic analysis (TA) [39] was used to generate in-depth insights from participants' discussions. A codebook-based TA approach, involving multiple coders (AMHM, TS, and NP), was used to enhance the credibility and consistency of the findings [40]. The primary analysis was conducted by 3 researchers, all of whom were experienced in conducting qualitative research and applying TA as an analytical method. Further analysis was conducted collaboratively by all authors of the study and validated by health care professionals to ensure the study addressed both theoretical and clinical aspects. The analysis followed Braun and Clarke's 6-phase framework, beginning with familiarization through repeated readings of the data to gain a comprehensive understanding. Subsequently, inductive and descriptive codes were applied, which were then grouped into preliminary themes. These initial themes were collaboratively reviewed and refined by the research team to ensure consensus about their accuracy and reliability. The processes of defining, naming the themes, and reporting the findings were conducted collectively.

To illustrate the development of the theme "safety and ease of use" using Braun and Clarke's 6-phase framework, we present an example from the analysis process. In the familiarization phase, repeated readings of transcripts revealed frequent concerns about physical safety, navigation, and psychological comfort in VR. Initial coding grouped statements about uncertainty, environmental awareness, and support needs under relevant codes, which were later collated into the broader theme. In the reviewing phase, we ensured this theme was coherent and differentiated from other related themes like "adaptability for stroke-related impairments," which focused more on task customization. "Safety and ease of use" emphasized risk minimization and user confidence. During theme definition and naming, we finalized it to reflect participants' emphasis on secure, intuitive, and supportive VR environments. Final reporting included selecting illustrative quotes to demonstrate how safety concerns informed both design recommendations and user engagement.

Rigor

Rigor was obtained through frequent debriefing and discussions among coders during data analysis and collaborative theme development within the research team. The qualitative data obtained from audio recordings of the workshops and individual interviews enabled triangulation and cross-references between findings [40], which strengthened the credibility of the results. The theoretical flexibility that TA offers was well-suited to accommodate the diverse perspectives and knowledge production of the multidisciplinary research team [41], which in this study included specialists in allied health, computer science, and design.

Ethical Considerations

The Norwegian National Research Ethics Committee for Medical and Health Research assessed the study and determined the need for ethical approval. They declared that this study's focus on health service research fell outside the scope of the Health Research Act § 2 and therefore did not require approval (ref 651236). Instead, this study was approved by the Norwegian Agency for Shared Services in Education and Research (ref 857865) to ensure compliance with privacy protection regulations. Participants received both oral and written information, and informed consent was obtained before the workshops and individual interviews. To protect the privacy of the participants, the study data were deidentified by using pseudonyms for all participants. Participants did not receive any financial or material compensation for their participation.

Results

Overview

Different participants with stroke, multidisciplinary health care professionals, VR developers, and researchers were present and participated in all workshops. There was a variety among the stroke individuals who were engaging in their first, second, or third workshop with the VR scenarios. Although engaging with avatars and objects in the café scenario was engaging, the stakeholders believed it did not target the rehabilitation principles. In the second workshop, this was amended by engaging in the kitchen and painting scenario. In the third workshop, these 2 prototypes were further developed by gamifying certain arm movements, targeting the affected arm in particular. The following results are based on the participants' needs and key challenges and benefits that were generated in workshops 1 and 2, and applied in workshop 3, which form the bases of this study, together with retrospective individual interviews with health care professionals. The findings generated five themes: (1) adaptability to accommodate stroke-related impairments in home rehabilitation, (2) safety and ease of use as fundamental in VR stroke rehabilitation, (3) goal orientation, (4) motivation, and (5) VR as a complementary tool.

Adaptability to Accommodate Stroke-Related Impairments in Home Rehabilitation

Participants viewed VR as a promising tool for poststroke home rehabilitation, offering interactive environments to aid motor recovery. However, they stressed that its usability depended on adapting to stroke-related impairments, such as accommodating both right- and left-handed hemiparetic users, as the first workshop primarily targeted right-hand users.

Customization was highlighted as crucial to address diverse needs, including issues like putting on VR glasses, gripping tasks, and aligning activities with rehabilitation principles. Participants stressed the importance of personalizing rehabilitation, as each person after a stroke faces unique challenges. Effective poststroke rehabilitation requires balancing engagement and endurance, as cognitive load and fatigue often limit people after a stroke. VR therapy demands motor and mental skills, making it vital to tailor intensity and duration to each individual's ability. It was noted that VR could overwhelm people after stroke, limiting its effectiveness.

> *When it comes to effectiveness, dosage and repetitions, VR will take more mental capacity because you use cognitive functions alongside motor function. Many people with stroke have fatigue; some can only do it for 10 minutes before their battery is low.* [Karla]

The potential for AI-powered VR solutions to enhance customization was discussed, with many participants emphasizing the rehabilitation goal of fostering independence and regaining agency. However, participants cautioned that VR was not suitable for everyone. Discrepancies between VR command techniques of gripping and actual abilities for people after stroke emerged during testing:

> *Therapist: You can also try to lift the kettle and the red bag there.*
>
> *Person after stroke, first-time user: I try with the left hand first.*
>
> *Person after stroke, first-time user: Just press through, sort of?*
>
> *VR developer: Yes. Use the whole thing. I had to practice a bit with gripping, because I thought I could grip with my fingers.*
>
> *VR developer: You have to use your whole hand, sort of, to grip.*
>
> *Therapist: Then you try to open your hand, and then... Like that, yes!*
>
> *Person after stroke, first-time user: Incredible.*

Participants also noted errors between VR scenarios and real-world tasks. One participant described challenges with gripping:

> *You can pick up things and put them in a drawer, but it's hard to grip them. Then it kind of falls and suddenly pops up again...You haven't managed to do what you intended...Maybe the tasks are too difficult or not well adapted.* [Jane]

Additional challenges included the difficulty of putting on heavy headsets for those with reduced arm function or configuring the VR system. One participant emphasized the need for customized adjustments to accommodate for stroke-related individual differences and goals:

> *I believe that those with severe cognitive impairments or difficulties with visual mapping might find VR too difficult initially. If we could personalize this, because one person's goal could be the opposite to others.* [Sue]

Participants recognized that VR required a certain level of cognitive function. Practicing new strategies in advanced tasks like kitchen training posed greater challenges than simpler tasks focused on errorless learning, such as dressing. Familiar settings were suggested to enhance engagement and skill transfer from virtual practice to real-world tasks. While cognitive impairments such as visual mapping and adjusting to new surroundings could cause difficulties, VR's potential to shift focus away from disability and reduce learned nonuse was noted. One participant explained:

> *When you enter that room, you slightly forget what is affected. And when you don't see that the arm isn't working, perhaps it will engage a bit more.* [Betty]

Kitchen training was widely regarded as a familiar and widely used element in rehabilitation. Participants noted that a VR kitchen could provide additional practice opportunities, as kitchens are a universally relevant and accessible setting for most people after stroke. A participant emphasized the importance of adjustable environments to accommodate individual needs.

"*You can put things high and low. You have the ability to increase or decrease the height of your entire kitchen, the cabinet, bench, whatever, so you can lower it or increase it for adjusting the difficulty level for your needs but also making it easier to use or potentially putting it to the exact same height as you have everything at home.* [John]

He also stressed the need for customizable difficulty levels, such as easy, medium, and hard, while acknowledging that further customizations were necessary to address individual differences. Many participants agreed that a one-size-fits-all approach would not work, but segmenting recovery stages into acute, subacute, and chronic phases, or tailoring tasks based on gripping ability, was proposed as a viable strategy.

Safety and Ease of Use as Fundamentals in VR Stroke Rehabilitation

Ensuring safety was a fundamental consideration in VR-based stroke rehabilitation. Many people after stroke face mobility impairments, balance issues, and cognitive challenges, increasing the risk of falls or disorientation in virtual environments. During the workshops, one health care professional highlighted their obligation to people after stroke:

> *We cannot tell them they can do it alone if we are not sure that it is safe.* [Jane]

Participants noted that a secure VR environment could encourage engagement and push physical limits, enhancing rehabilitation progress. However, some participants warned that VR might create a false sense of safety. As one participant with stroke explained:

> *Now I can't see where my legs are or anything. It's an uncertainty. I feel like going over to hold onto the table, but that won't work. Now I'm trying to do as I do at home and park the walker next to the kitchen door.* [Person after stroke, second-time user]

The transition from institutional rehabilitation to home-based therapy was identified as a critical phase, often accompanied by anxiety and uncertainty. One participant noted that VR could reduce anxiety and prevent early readmission to health care facilities:

> *We have to make sure they go home without as much anxiety, because when they do, they might just return early to the healthcare system due to a fall or stress from something minor. With the help of virtual reality, we could avoid this because we've tested it already at "home", but at the institution.* [Tom]

Participants emphasized the need to carefully consider the difference between VR tasks and real-world activities. He further explained:

> *That's going to be challenging, since there is a clear difference between how difficult something will be in VR versus in the real world, like cutting vegetables with a knife.* [Tom]

Some participants noted that successfully completing tasks in VR but failing in real life could lead to disappointment, especially when tasks required a higher degree of task complexity.

One highlighted the importance of minimizing unnecessary movement in home-based VR rehabilitation:

> *If someone is going to use it alone, I would limit tasks where you have to walk far, because I wouldn't consider that safe.* [Jane]

Similarly, Kate pointed out that while real-life practice is ideal, VR enables people after stroke to safely train in everyday activities like walking, empowering them to actively engage in their recovery.

Beyond physical safety, participants also discussed the psychological benefits of VR, particularly its ability to reduce the fear of failure, which often prevents people after stroke from participating in rehabilitation activities. Ben underscored the importance of structured support for VR in both institutional and home settings:

> *Using it in our hospital department would be great because they can use it in their rooms with instructions to be seated during activities. And it can be used in the later afternoon when they're done with daily [therapy-assisted] rehabilitation training. At home, they would be instructed to sit in a chair or a sofa with a space around them in case they get excited and get up.* [Ben]

The focus on sitting exercises typically placed occupational therapists in charge of VR rehabilitation due to traditional task distribution. However, health care professionals raised concerns about the safety and feasibility of using the VR system without therapist assistance, questioning whether users could operate and administer the system without assistance. Technological advancement, such as developing hybrid and mixed reality VR rooms, were proposed as potential solutions to address safety concerns. One participant with stroke explained:

> *I would need the ability to sense the outline of it [the walker], because it is safer when I know where it is. Otherwise, I could have practiced in the real kitchen, so it was fairly similar.* [Person after stroke, second-time user]

Furthermore, the presence of a caregiver or therapist was seen as essential for building user confidence in using VR, as indicated by Maud:

> *It's very important to have someone who watches what you and [make sure] you do it carefully. If you fall, there's someone with you, and you feel secure. Our people with stroke have belts on, and the physio holds the belts from the back. Maybe these people with stroke can't walk without a stick, but with that support, they feel very secure.* [Maud]

One proposed solution, which came up during workshops, was to differentiate between safer, low-scale tasks for home use and more challenging tasks, such as standing or increased reaching activities, to be conducted during follow-up sessions at the hospital.

Goal Orientation

The participants highlighted that a goal-oriented approach was essential in poststroke rehabilitation, as it helped people stay motivated and engaged after stroke. VR was perceived as a platform to structure rehabilitation programs around individualized goals, allowing people after stroke to track progress and experience achievement. The participants found that incorporating meaningful, measurable objectives encouraged consistent participation and reinforced the benefits of repetitive practice. Observing progress and perceiving tangible improvements increased commitment to rehabilitation exercises. Several participants stressed the importance of person-centered VR rehabilitation, where goals align with daily needs and personal aspirations. Karla noted:

> *The person with stroke's goal is everything in the rehabilitation. If someone wants to dress independently, we have to focus on that. The person with stroke is in the driver's seat.* [Karla]

Participants emphasized that VR technology should focus on customization in goal setting to enable people after stroke to work on relevant activities, like regaining independence in dressing, cooking, or mobility. To maximize effectiveness, VR interventions needed a clear focus.

> *I think it's really important if you use this in a rehabilitation program, you need to specify your target. Is it cognition, arm movement, or balance?* [Tom]

However, limitations were noted where VR could not adapt to all needs. During the workshops, a health care professional explained:

> *Some patients can only use this movement—just this—but others need to grip something. Then, the patients need resistance, and this [VR] won't improve their function like the real activity. If they are going to move the paint, there is some weight in the paint, and you dont get transferability. To see improvement at home, they must do it similarly.* [Physiotherapist, Workshop 3]

Participants suggested that providing multiple training options would enhance adherence and long-term commitment to therapy. Effective VR rehabilitation required targeting more tasks that mirror real-life activities to develop skills essential for independent living after discharge, such as putting up a painting, washing the dishes, or putting on clothes. It was important that the chores or activities used in rehabilitation served a specific therapeutic purpose. Kate pointed out:

> *An activity like painting, you don't want to be a better painter, but you want your arm to get better. That's the method.* [Kate]

From a therapist's perspective, this highlights how everyday tasks can be repurposed as meaningful rehabilitation goals, where the focus is on physical recovery rather than mastering the activity itself. A blended approach, where people after stroke first practiced activities in VR and then transferred skills to real-life settings, was suggested to reinforce motor learning and recovery. Repeating engagement in virtual tasks before real-life performance helped internalize movements, build confidence, and strengthen neural connections essential for motor function.

However, not all participants' goals were covered by VR prototypes. One participant with stroke noted:

> *There is no gait training, at least not intentionally. I would have to be on a treadmill, walking into the woods. This is for the arm, though.* [Person after stroke, second-time user]

Tracking progress in VR was seen as a powerful motivator for people after stroke. Participants explained that observing measurable improvements encouraged continued training and inspired confidence to tackle more demanding tasks. They emphasized that VR should serve as a stepping stone, bridging the gap between virtual practice and real-life application. However, there was a need for a clear focus on whether VR training aimed for improvement or sustainability. Gamification elements, like points or rewards, were considered ineffective unless the real-world relevance of the tasks was clear for people after stroke:

> *It's not necessarily the case that you get motivated by getting extra points, if you don't quite understand why you're doing this.* [Jane]

Ensuring VR-based rehabilitation included familiar, everyday tasks enabled explicit goals in a more engaging, functional, and transferable home environment. Therefore, a variety of VR rehabilitation activities, such as hanging clothes in the laundry room or other meaningful daily activities, were considered valuable.

Motivation

The participants emphasized that motivation played a crucial role in the effectiveness of VR rehabilitation for people post stroke, ensuring long-term engagement with their rehabilitation program. Tom noted the importance of regular follow-up sessions for maintaining motivation.

> *We can see when they were very motivated here. They go home, and there are only challenges, no solutions. But when they return for follow-up sessions, they're extremely motivated.* [Tom]

Motivated individuals were more likely to push themselves and put in the effort required to achieve their rehabilitation goals, potentially leading to better motor outcomes and quality of life. People with stroke noted that the VR scenarios stimulated learning and movement, enabling them to train without being consciously aware of it.

> *Person after stroke, third-time user: Now I have used the entire room, I see.*
>
> *Therapist: Now we have followed you around the entire room.*
>
> *Person after stroke, third-time user: Today's workout.*
>
> *VR developer: Shall we see how many points there have been. It sounded like it was beeping a lot.*

Karla highlighted how competition could drive people after stroke to improve their performance.

> *It's motivating when it's like a competition. You did this last time, now try to reach this one for the progression of function.* [Karla]

Setting and achieving progressively challenging goals allowed people after stroke to experience significant functional gains, reinforcing their overall commitment to recovery and reaching rehabilitation outcomes. One participant noted the extrinsic motivation provided by VR and its effects on the brain.

> *Because when you do that, your brain tells you it's very useful, because you get points.* [Person after stroke, second-time user]

While the rehabilitation process could be challenging and frustrating, motivation was seen as a driver to overcome the physical and emotional barriers, persisting through difficulties, and maintaining a positive attitude toward recovery. Many participants emphasized the role of human support in maintaining motivation, as Ben explained:

> *If you're not motivated or very depressed, your life is over. We won't do anything with you. You need humans to train with you.* [Ben]

Health care professionals' emotional and psychological encouragement was regarded as crucial in helping people after stroke stay engaged. Motivated individuals were more likely to communicate their preferences and needs, allowing for a more personalized and active rehabilitation experience.

In addition, motivation was closely linked to psychological well-being. Engaging in meaningful and enjoyable activities improved mood, reduced anxiety, and enhanced overall mental health, which was crucial for recovery. Sue underscored the role of curiosity in maintaining motivation.

> *It's more like an opportunity to explore and be curious. If you manage to get curious, you have less capability to be afraid or be sceptical.* [Sue]

Building confidence was another key factor in motivating people after stroke to embrace new challenges and persist with VR. One suggested strategy to foster confidence was visual progress tracking.

> *If you could implement that, for example, in painting stations, storing your paintings so you could see at the beginning I was able to make a line that was so big, but after two weeks I'm able to make a line that's this big, could be a motivational tool.* [Sue]

Seeing tangible improvements over time reinforced the belief of people after stroke in their abilities, motivating continued rehabilitation efforts. Jane emphasized the importance of simplicity and ease of use for maintaining motivation.

> *For people with stroke to use it themselves, they must have high motivation. It must be simple enough to master quickly, not something they struggle with and give up in the end. It must be meaningful and beneficial for them.* [Jane]

Others suggested integrating VR with existing digital training programs, such as Exorlive, to offer a self-training option to therapy-assisted activity training. Ben highlighted the difference in outcomes between motivated and unmotivated individuals.

> *If you're a person who wants to do better, you have a very clear goal. You just do that with the help of others or sometime without any help. But if you're a person like, oh, now this is my life, I'm going to live*

> *with it, and you're not motivated to do anything, you might just sit there and let it go day by day.* [Ben]

Those with clear goals and motivation were likely to strive for improvement, while those without motivation might become complacent.

Finally, John emphasized the importance of creating varied and personalized scenarios to maintain motivation, stating:

> *To motivate people, it would be more interesting to create different environments, such as a bedroom where you can dress yourself or a garage where you can tinker and build with your hands.* [John]

By providing scenarios aligned with the specific goals and interests of people after strokes, the rehabilitation process becomes more rewarding and engaging, fostering long-term motivation.

VR as a Complementary Tool

VR was seen as a valuable complement to stroke rehabilitation, enhancing training intensity and offering structured exercises in both clinical and home settings. While not a replacement for traditional therapy, it effectively supported recovery by supplementing existing approaches. Given the limited window for optimal recovery, its role in maximizing progress was emphasized. Mary highlighted VR's contribution to intensive training.

> *Our institution specializes in giving intensive training. I don't think VR could match that, but as an additional supplement it would be amazing. Together we could help people with stroke recover more quickly. Especially for people with stroke, there is a time period to get as good as they can get.* [Mary]

This highlights that VR should complement, not replace, therapist-led rehabilitation by enhancing therapy intensity during the critical recovery phase. Additionally, participants proposed implementing validated clinical measurements, goal assessment scales, or kinematic data to monitor outcomes and progress within the VR system during the workshops. Betty noted the alignment between the painting scenario and existing cognitive tests, such as MoCA, suggesting VR's potential for cognitive diagnostics in addition to general executive functions:

> *Because we have a test where we follow a line with a pen from number to number, or from numbers to letters. That kind of cognitive test. So, both for cognition and physical aspects, it [VR] is, if I may say so, very useful.* [Betty]

Individuals with stroke face unique challenges, requiring personalized rehabilitation plans tailored to their specific needs. VR adaptability offers a promising tool for delivering customized training that complements traditional therapy, ensuring alignment with each person's circumstances. VR was seen as a tool to bridge this gap by providing structured exercises that encourage movement and engagement.

> *It's another intervention in addition to what we do. I think it would increase the activity for the people with stroke to use their arm more than they otherwise would, if they have such an aid in addition to a therapist.* [Jane]

During workshops, VR was seen as a controlled environment where people after stroke could practice real-life tasks safely before transitioning to real-world applications. Jane highlighted this advantage:

> *We've discussed it being a supplement to training and in a safer setting for the people with stroke. It will not be instead of, but it may increase the activity between each session. Someone has concrete things to practice until next time, then practice in a real kitchen situation when someone is present to secure the situation.* [Jane]

Combining home-based VR training with in-person therapy was considered an ideal model for stroke rehabilitation. Participants, like Kate, emphasized that while VR enabled frequent, self-paced training at home, regular sessions with a therapist were essential to maintain motivation. This hybrid approach demonstrated VR's potential to extend rehabilitation beyond clinical settings while preserving professional oversight. Balancing self-practice with therapist support was considered crucial. However, as Betty noted, the attitude and behavior of staff introducing VR could significantly influence patient engagement.

> *Let's say the therapist has zero interest in VR. That could easily rub off on the patient. (..) I think it makes a big difference who introduces it to the patients.* [Betty]

As stroke rehabilitation requires extensive repetition to regain lost motor functions, VR was seen as an efficient way to deliver high-volume training, ensuring people after stroke get the practice needed for meaningful improvement. One participant explicitly noted they lacked solutions targeting volume training:

> *The advantage of VR is that you can do high-volume training efficiently. In rehabilitation of people with stroke, you need to spend hundreds of training hours to achieve some really big goals.* [John]

VR was seen as a possible way to make repetitive training more accessible and help people commit to the long-term rehabilitation process necessary for significant recovery. While VR was valuable, it was perceived as a stepping stone for real-life training. During the second workshop, an occupational therapist participant emphasized the limitations of VR in meeting diverse patient needs, noting that some individuals require functional resistance and weight-bearing activities to achieve transferability to real-life tasks:

> *Most people understand that VR is a supplementary tool along the way that bridges the gap, making it easier, but they know they need real life [activities] because that's where they are anyway.*

These results show that VR should not replace traditional therapy but serve as a tool to facilitate a smoother transition to functional activities.

Summary of the Key Findings

The integration of VR into stroke home rehabilitation presents both significant challenges and promising benefits. One of the primary challenges lies in the adaptability of VR systems for individuals with hemiparesis, as these systems must accommodate both right- and left-handed users. This requires considerable technical resources and expertise to ensure that the technology can be effectively tailored to the diverse needs of people after stroke. Additionally, customization remains a complex barrier, as VR solutions must be capable of addressing a wide range of stroke-related impairments and aligning with specific rehabilitation goals, which can vary greatly from person to person. Another critical concern is the balance between engagement and endurance. While VR has the potential to make rehabilitation more stimulating, it is essential to calibrate the intensity and duration of exercises to prevent cognitive overload and fatigue, which are common among individuals with stroke. Usability also poses a significant hurdle; many people after stroke experience limited motor function, which can make it difficult to interact with VR headsets and systems. Furthermore, challenges in visual mapping and task comprehension can hinder the effectiveness of VR-based interventions.

Despite these challenges, VR offers several compelling benefits for enhancing stroke rehabilitation at home. The integration of AI into VR platforms can enable highly personalized rehabilitation experiences, adapting in real time to the user's progress and specific needs. This level of customization can significantly enhance the efficacy of home-based therapy. Moreover, VR has the potential to increase motivation and engagement through gamification, progress tracking, and the simulation of meaningful, real-life tasks that resonate with users. VR also provides a controlled and safe environment for practicing daily activities, which can help reduce anxiety and better prepare people after stroke for real-world challenges. As a supplementary tool, VR can enhance traditional rehabilitation by increasing the intensity and volume of therapy. This aligns with 10 key principles of experience-dependent neuroplasticity, where factors such as specificity, high-repetition, intensity of stimulation, timing, task-specific training, auditory stimuli, and behavioral experience are essential for optimizing recovery outcomes after brain damage [42]. Textbox 1 provides an overview of the main challenges and potential benefits associated with using virtual reality in poststroke rehabilitation. It highlights critical considerations for implementation and the opportunities VR offers to enhance recovery outcomes.

Textbox 1. Summary of the key challenges and benefits in the use of VR for rehabilitation of people after stroke.

Key challenges in virtual reality (VR) usage for people after stroke rehabilitation

- Adaptability for hemiparesis
- VR systems need to accommodate both right- and left-handed users, which requires significant resources and skill to adapt to various people after stroke needs.
- Customization barriers
- VR technology must address individual stroke-related impairments and specific rehabilitation goals, which can be complex and demanding.
- Balance between engagement and endurance
- Tailoring the intensity and duration of VR exercises to avoid overwhelming people after strokes is crucial, especially considering cognitive load and fatigue.
- Usability concerns
- VR headsets and systems can be challenging for people after strokes with limited motor function, and difficulties in visual mapping and understanding tasks.

Key benefits in VR usage for people after stroke rehabilitation

- Enhanced customization with artificial intelligence
- Artificial intelligence–powered VR solutions can provide personalized rehabilitation experiences, address individual needs, and enhance home-based therapy.
- Motivation and engagement
- VR can make rehabilitation more engaging through gamification, tracking progress, and providing meaningful, real-life tasks that motivate people after strokes.
- Controlled and safe environment
- VR offers a controlled environment for practicing real-life tasks safely, reducing anxiety and preparing people with strokes for real-world challenges.
- Supplement to intensive training
- As a complementary tool, VR can increase therapy intensity, providing high-volume, repetitive training efficiently to maximize recovery outcomes.

Discussion

Principal Findings

This study explored the challenges and benefits of co-designing and using VR to support home-based poststroke rehabilitation, with a particular emphasis on user involvement and experience. The TA revealed 5 key themes that offer insight into how VR can be effectively integrated into rehabilitation practices. Adaptability emerged as a central concern, highlighting the need for VR systems to accommodate a wide range of stroke-related impairments, including motor, cognitive, and sensory limitations. Participants emphasized the importance of customizable interfaces and adjustable task difficulty to ensure accessibility and relevance. Safety and ease of use were also critical, with users expressing a preference for intuitive designs that minimize physical strain and cognitive overload, especially in unsupervised home settings. The theme of goal orientation underscored the value of VR in facilitating task-specific training that aligns with personal rehabilitation goals, thereby enhancing the perceived relevance and purpose of exercises. Motivation was closely linked to the immersive and interactive nature of VR, which participants found to be more engaging than traditional methods, potentially supporting persistent participation over time. Finally, the role of VR as a complementary tool reflected a shared understanding that VR should not replace conventional therapy, but rather augment it, offering additional benefits for practice and reinforcement in a flexible, home-based format.

The study has revealed that one of the primary benefits lies in the adaptability and individualization of VR environments to meet the diverse needs of people after stroke. As noted in the study, many people after stroke experience hemiparesis, necessitating VR systems that accommodate both right- and left-affected users. It was noted that the final prototype involving the kitchen scenario factored this in by rewarding participants for using their affected arm. The ability to individualize VR experiences to address specific activities impaired by stroke, such as initially setting up the VR system and further conducting activities, such as reaching and gripping objects, is critical for effective rehabilitation. With appropriate individualization, VR systems have demonstrated significant effectiveness in improving various aspects of stroke rehabilitation, including upper limb function, cognitive abilities, and balance [25,43,44].

However, to bridge the gaps between the recovery process in stroke rehabilitation, technology, and clinical practice, an awareness, exploration, experimentation, and evaluation framework ensuring the solutions are engaging, accessible, accountable, and adaptable is advocated [45]. A common challenge with commercially available gaming technology, such as VR, is that even though it provides immersive, engaging, and tailored therapy (accessibility and engagement), it is not easily adapted for people after stroke, especially those experiencing limited movement or cognition. Based on our assessment, we argue that the complexity of the devices makes them difficult to adjust, posing a challenge for people after stroke to secure them properly and customize them to their individual needs. Although customization was frequently proposed to meet the individual needs of users, this highlights a critical tension. In practice, effective customization demands considerable resources to adequately and efficiently respond to the diverse and complex needs of individuals living with stroke. As indicated by one study, reducing the complexity to one working application on the goggles can limit potential errors [23]. In this study, some individuals required additional assistance in adjusting the headset or coping with cognitive impairments. There is a paradox between the ideal of customization and the reality of implementing it effectively. Achieving this may involve a balance between providing customized opportunities and still making it feasible to implement and ensure its ease of use by categorizing similar users into groups.

VR has the benefit of providing multiple training environments that target individuals by engaging them in meaningful and enjoyable activities, as particularly found in the kitchen scenario. Functional exercises, such as washing dishes, dressing, and brushing teeth, allow people after stroke to develop skills essential for independent living after discharge [46]. This aligns with one study [47], in which people after stroke found different activities in everyday life easier to perform since starting their VR training, such as opening a drawer or applying toothpaste to a toothbrush. Enriched environments are essential in leveraging VR for stroke rehabilitation, as they offer diverse and stimulating approaches to problem-solving and skill development [48]. Some overarching principles that guide an enriched environment approach are complexity, variety, and novelty of the environment as well as targeting underlying needs [49]. In our study, the VR scenarios targeted all principles, but ideas for scaffolding complexity and variety tailored for each individual were proposed.

By simulating real-life scenarios and interactive challenges, VR provides people after stroke with engaging and repetitive practice, which is crucial for neuroplasticity and motor recovery [50]. Additionally, the adaptability of VR allows therapists to modify the learning environment to match individual needs, for example, by adjusting the levels of difficulty, thereby fostering a more immersive and intensive experience that enhances engagement and accelerates rehabilitation progress. This flexibility not only improves functional recovery but also increases motivation and adherence to therapy, making VR a promising tool in modern rehabilitation strategies [51]. In a study by Gustavsson et al [47], VR provided a feeling of being in a different world where users felt they were reaching higher and moving faster. However, VR environments can deceive the mind into believing they are a real-life environment. This illusion is a cause for concern, as the immersive nature of VR can make the brain perceive the simulated surroundings as authentic, thereby inducing a false sense of safety and achievement [52].

Given the challenging nature of rehabilitation for individuals who have experienced a stroke, VR plays a critical role in motivating them by ensuring their active participation and long-term effort in the recovery process. Engaged people after stroke are more likely to adhere to their therapy schedules and participate actively in their rehabilitation exercises. Rehabilitation can be challenging and frustrating, and motivation helps people after stroke overcome these barriers, persist through difficulties, and maintain a positive attitude toward their recovery [25]. It drives users to push through the physical and emotional challenges of rehabilitation, helping them to stay committed to their therapy even when progress seems slow or when they encounter setbacks. By fostering a sense of achievement and progress, motivation can transform the rehabilitation experience from a daunting task into a rewarding journey.

However, maintaining motivation can be difficult, especially for people after stroke facing mental health challenges, such as depression or anxiety. These conditions can significantly impact their willingness to engage in rehabilitation activities. While we saw that the use of gamification elements, such as points and rewards, served as incentives to motivate people after stroke, they work only if people after stroke understand the real-world relevance of their tasks. Without this understanding, the motivational impact of these gamification elements may be significantly diminished. This highlights the need to design VR programs that closely mimic real-world scenarios to maximize their effectiveness.

Another benefit of VR is that it creates a controlled and secure environment that encourages people after stroke to engage more actively in rehabilitation by reducing the fear of injury, thereby promoting better recovery outcomes when they feel safe enough to push their physical limits. However, this is contingent on structured supervision and support, which are essential to mitigating risks and enhancing user confidence, particularly given the mobility impairments, balance issues, and cognitive challenges faced by people after stroke, which increase the risk of falls or disorientation [25,44]. The importance of therapist involvement in tutorials and adjusting difficulty levels when engaging with VR is highlighted in multiple studies [23,47]. This is particularly important for individuals recovering from stroke, as prolonged exposure to VR may induce motion sickness or dizziness. Additionally, those with impairments in executive functioning and problem-solving challenges observed during our workshops may find it especially difficult to navigate and adapt to such technologies. To maximize the benefits of VR therapy, careful system design, real-time monitoring, and gradual progression in exercise intensity are essential to ensuring a safe and effective rehabilitation experience [53].

While support and guidance in the use of VR emerged as indispensable, this could present challenges in terms of manpower due to the decreasing number of health care professionals. The presence of therapists significantly motivated many individuals with stroke to continue engaging with VR activities. However, this reliance on external motivation could become problematic once they are discharged home, where self-motivation is essential for ongoing rehabilitation [54]. It is not feasible to maintain this level of support when individuals with stroke transition to their home setting. Many people with brain injuries are dependent on help from their families when using digital technologies, such as VR, at home [23,54].

The use of VR presented an advantage in facilitating a goal-oriented approach by enhancing consistent participation and reinforcing the benefits of repetitive practice. The application of motor learning principles in VR design is advocated to enhance rehabilitation outcomes [25]. This includes progressively challenging tasks, real-time performance feedback, and multimodal feedback [25,43]. The development process through the workshops led to improvements in VR content in line with standardized stroke guidelines, such as goal- and task-specific activities and quantifying activities. However, it emerged from the study that the challenge lies in ensuring that VR interventions are carefully designed with clear and specific goals, whether to improve cognitive function, enhance arm movement, or restore balance. It is important that the rehabilitation goals are aligned with the daily needs and personal aspirations of the people after stroke. This requires a deep understanding of each person’s lifestyle, preferences, and long-term objectives. Achieving this alignment is challenging, as the goals can change over time, requiring continuous assessment and adjustment of the VR activities to ensure they remain relevant and motivating.

Despite the positive effects associated with measuring progress in VR, the process can be difficult [55]. This difficulty arises from the need to ensure that the metrics used in VR accurately reflect real-world abilities and improvements. While many VR games are primarily designed for entertainment rather than rehabilitation, the importance of involving users in the co-design of VR applications is increasingly recognized [25].

The use of VR in this study revealed that it is beneficial as a complementary tool in stroke rehabilitation, enhancing user engagement, increasing training intensity, and providing structured exercises in both clinical and home environments. While VR cannot replace traditional rehabilitation methods, it can serve as a powerful addition that supplements existing therapy approaches. Previous studies suggest that integrating specific VR technologies with traditional rehabilitation approaches can lead to greater improvements in motor function and activity levels in people after stroke compared with using conventional rehabilitation methods alone [56-58]. For example, VR can provide repetitive and varied exercises that target specific motor skills, helping people after stroke to practice and refine their movements in a controlled environment, thereby freeing therapists from the mundane tasks of high-volume training.

A potential benefit could also be, as found in a study [59], that VR demonstrated significantly greater therapeutic effects than conventional training in improving upper limb function, as assessed by the Fugl-Meyer and Action Arm Research Tests [32]. Following the stroke trajectory, VR could be implemented during the hospital stay, as a prepractice and supplement to real-life activities. In the home setting, it would need to be scaled down to accommodate safety precautions, but has the possibility to be further scaffolded when returning for follow-up at the hospital in cooperation with health care professionals. A significant challenge in using VR for rehabilitation is ensuring its effective integration with traditional therapy while maintaining professional oversight as rehabilitation extends beyond clinical settings.

Successful integration is essential, as it necessitates a seamless combination of VR-based and conventional approaches to optimize therapeutic outcomes. Professional oversight remains crucial for monitoring patient progress, adjusting interventions, and providing ongoing support tasks that become increasingly difficult to maintain once people transition to home-based rehabilitation after stroke.

Implications for Practice

Our findings suggest that VR could be a valuable supplement to traditional rehabilitation for specific stroke groups, but it cannot replace the essential contact with the therapist, clinical reasoning, or the assessment of activity limitations and opportunities. We propose that VR could be particularly beneficial for a younger population with stroke who lack sufficient access to home-based rehabilitation and struggle with motivation for self-training. Additionally, VR may help maintain function in individuals in the chronic phase after a stroke, where repetition and high-quality training are crucial to prevent functional decline. However, it is important to acknowledge that VR cannot fully replicate training in real-life, everyday activities. For VR solutions to be effective, they must be carefully tailored to the individual's functional level and ease of use. This tailoring process, however, can be resource-intensive in an already strained health care system.

In light of this, we raise important questions about who will truly benefit from VR, given the wide range of outcomes following a stroke, and whether, in some cases, it may be more appropriate to prioritize traditional therapy. Finally, we emphasize the importance of training for relatives and caregivers, as their role may be critical to the successful use of VR at home.

Limitations and Future Work

The limitations of this study primarily stem from the use of a qualitative single case study design [27], which provides an in-depth understanding of the real-life phenomenon but may limit the generalizability of the findings across broader contexts. The reliance on rapid prototyping and TA, while effective for capturing detailed insights into VR usage in poststroke rehabilitation, may not fully address the variability of experiences and needs among a wider population of people after stroke. Also, the rapid evolution of VR technology means that some of the results will soon be outdated, limiting the applicability of the findings to current and future developments in the field [55,60].

Future work should aim to expand the scope of research by incorporating multiple case studies across diverse settings to enhance the generalizability of the results. This could involve a larger and more varied participant pool to capture a broader range of experiences and needs. Future work should focus on developing integrated, individualized therapy solutions that seamlessly combine arm therapy, cognitive therapy, and mental support within a single solution, addressing the multifaceted rehabilitation needs, such as cognitive needs or physical therapy needs, of individuals with stroke as described in Gkintoni et al [61].

This requires implementing design methodologies that facilitate co-creation with diverse stakeholders, including health care professionals, VR developers, and people after stroke with varying needs, ensuring that the solutions are both inclusive and effective. Efforts should also be directed toward creating a smooth transition from hospital settings to home environments for people after stroke, as described by Pourliaka et al [62], using VR to extend rehabilitation beyond clinical settings while maintaining professional oversight.

Furthermore, the development of AI-powered solutions that account for the personalized goals, needs, and situations of people after stroke is crucial, enabling tailored interventions that adapt to individual progress and challenges. For instance, the system could select the most appropriate exercises based on how individuals perform each exercise [63]. One option is to incorporate validated assessments, such as the short-form of MoCA [31,64], to evaluate patient situations accurately and guide therapy adjustments.

Designing solutions that not only excite individuals with stroke, but also genuinely facilitate improvements in their therapy goals and daily life challenges is vital. This involves crafting engaging VR environments that motivate consistent participation while ensuring real-world relevance and measurable benefits in rehabilitation outcomes. Further exploration of potential scenarios that can adapt to the individual requirements of home rehabilitation over time is also essential.

Additionally, more validation studies, for example, pilot trials or feasibility studies to evaluate the preliminary effects and randomized controlled trials to evaluate actual VR solution outcomes, are urgently needed [54]. In addition, standardized usability and comfort measures, such as the System Usability Scale and the Virtual Reality Sickness Questionnaire, were not used. Their inclusion could have offered more structured and quantifiable assessments of user experience, particularly regarding interface usability and potential adverse effects like motion sickness. Future work should consider integrating these validated instruments to complement qualitative feedback and provide a more comprehensive evaluation of user interaction and comfort.

Conclusion

This study underscores the transformative potential of VR in advancing stroke rehabilitation. VR provides a dynamic and interactive platform that can be customized to support the individualized trajectories of recovery, accommodating the physical limitations, cognitive profile, and personal goals of people after stroke. Its core strength lies in its adaptability by offering immersive simulations of real-world tasks that foster motivation and engagement through meaningful and goal-directed experiences. However, the successful application of VR in rehabilitation requires more than technological advancement. It needs to be integrated thoughtfully with established rehabilitation principles, including task-specific training, measurable and ecologically valid outcomes, as well as alignment with the evolving needs of individuals. A critical insight from this study is the necessity of embedding clear therapeutic intent into VR interventions. Whether the objective is motor recovery, cognitive enhancement, or balance restoration, each activity should serve a purpose that is clinically relevant and meaningful to the user's daily life. This demands a comprehensive understanding of individual motivations, challenges, and aspirations through involving people after stroke in the design process. It is through participatory co-design that developers can ensure that VR experiences are both engaging and clinically appropriate. As an emerging technology, VR should not be considered a replacement for traditional therapy, but rather a powerful adjunct that can increase training intensity, reduce the burden on therapists, and extend rehabilitation into home settings, making therapy more accessible and continuous. When thoughtfully designed and implemented, VR can bridge the gap between structured clinical care and independent recovery, providing consistent support throughout the rehabilitation journey.

Funding

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

Data Availability

The dataset is derived from interviews in Norwegian and is not publicly available due to consent restrictions and data collection approval. Parts of the data can be made available from the corresponding author upon request.

Authors' Contributions

Conceptualization: NP, MAP
Methodology: NP, MAP, TS
Investigation: TS, AMHM, NP, IKHR, KGSO, NSM, ÅB
Data curation: TS, AMHM
Formal analysis: All authors
Project administration: TS, NP, MAP
Resources: IKHR, KGSO, NSM, ÅB
Validation: All authors
Visualization: NP, AMHM
Writing – original draft: TS, AMHM
Writing – review and editing: TS, AMHM, IKHR, KGSO, NSM, ÅB, NP, MAP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview and workshop guide.
[DOCX File , 98 KB - humanfactors_v13i1e78798_app1.docx]

References

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
2. Fugazzaro S, Denti M, Accogli M, Costi S, Pagliacci D, Calugi S, On Behalf Of Look After Yourself Project. Self-management in stroke survivors: development and implementation of the Look after Yourself (LAY) intervention. Int J Environ Res Public Health 2021 May 31;18(11):1-14 [FREE Full text] [doi: 10.3390/ijerph18115925] [Medline: 34072998]
3. Saini V, Guada L, Yavagal DR. Global epidemiology of stroke and access to acute ischemic stroke interventions. Neurology 2021;97(20 Suppl 2):S6-S16. [doi: 10.1212/WNL.0000000000012781] [Medline: 34785599]
4. Kwakkel G, Stinear C, Essers B, Munoz-Novoa M, Branscheidt M, Cabanas-Valdés R, et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. Eur Stroke J 2023 Dec;8(4):880-894 [FREE Full text] [doi: 10.1177/23969873231191304] [Medline: 37548025]
5. Bui J, Luauté J, Farnè A. Enhancing upper limb rehabilitation of stroke patients with virtual reality: a mini review. Front. Virtual Real 2021 Nov 8;2:595771. [doi: 10.3389/frvir.2021.595771]
6. Dias P, Silva R, Amorim P, Lains J, Roque E, Pereira ISF, et al. Using virtual reality to increase motivation in poststroke rehabilitation. IEEE Comput. Grap. Appl 2019;39(1):64-70. [doi: 10.1109/mcg.2018.2875630]
7. Sahely A, Giles D, Sintler C, Soundy A, Rosewilliam S. Self-management interventions to improve mobility after stroke: an integrative review. Disabil Rehabil 2023 Jan;45(1):9-26 [FREE Full text] [doi: 10.1080/09638288.2022.2028019] [Medline: 35068313]

8. Meyer S, Verheyden G, Brinkmann N, Dejaeger E, De Weerdt W, Feys H, et al. Functional and motor outcome 5 years after stroke is equivalent to outcome at 2 months: follow-up of the collaborative evaluation of rehabilitation in stroke across Europe. Stroke 2015;46(6):1613-1619. [doi: 10.1161/strokeaha.115.009421]
9. Lieshout ECCV, van de Port IG, Dijkhuizen RM, Visser-Meily JMA. Does upper limb strength play a prominent role in health-related quality of life in stroke patients discharged from inpatient rehabilitation? Top Stroke Rehabil 2020 Oct;27(7):525-533 [FREE Full text] [doi: 10.1080/10749357.2020.1738662] [Medline: 32148183]
10. Schnabel S, van Wijck F, Kidd L. Supporting stroke survivors to meet their personal rehabilitation needs in community-based arm rehabilitation: development of initial programme theories to explore what may work for whom, how and under what circumstances. Front Neurol 2023;14:1089547 [FREE Full text] [doi: 10.3389/fneur.2023.1089547] [Medline: 37332992]
11. Mekbib DB, Han J, Zhang L, Fang S, Jiang H, Zhu J, et al. Virtual reality therapy for upper limb rehabilitation in patients with stroke: a meta-analysis of randomized clinical trials. Brain Inj 2020 Mar 20;34(4):456-465. [doi: 10.1080/02699052.2020.1725126] [Medline: 32064964]
12. Domínguez-Téllez P, Moral-Muñoz JA, Salazar A, Casado-Fernández E, Lucena-Antón D. Game-based virtual reality interventions to improve upper limb motor function and quality of life after stroke: systematic review and meta-analysis. Games Health J 2020 Feb;9(1):1-10 [FREE Full text] [doi: 10.1089/g4h.2019.0043] [Medline: 32027185]
13. Henderson A, Korner-Bitensky N, Levin M. Virtual reality in stroke rehabilitation: a systematic review of its effectiveness for upper limb motor recovery. Top Stroke Rehabil 2007;14(2):52-61. [doi: 10.1310/tsr1402-52] [Medline: 17517575]
14. Kang D, Park J, Eun S. Home-based virtual reality exergame program after stroke rehabilitation for patients with stroke: a study protocol for a multicenter, randomized controlled trial. Life (Basel) 2023;13(12):2256 [FREE Full text] [doi: 10.3390/life13122256] [Medline: 38137857]
15. Clark WE, Sivan M, O'Connor RJ. Evaluating the use of robotic and virtual reality rehabilitation technologies to improve function in stroke survivors: a narrative review. J Rehabil Assist Technol Eng 2019;6:2055668319863557 [FREE Full text] [doi: 10.1177/2055668319863557] [Medline: 31763052]
16. Park D, Lee D, Lee K, Lee G. Effects of virtual reality training using Xbox Kinect on motor function in stroke survivors: a preliminary study. J Stroke Cerebrovasc Dis 2017;26(10):2313-2319. [doi: 10.1016/j.jstrokecerebrovasdis.2017.05.019] [Medline: 28606661]
17. Sin H, Lee G. Additional virtual reality training using Xbox Kinect in stroke survivors with hemiplegia. Am J Phys Med Rehabil 2013;92(10):871-880. [doi: 10.1097/PHM.0b013e3182a38e40] [Medline: 24051993]
18. Moan ME, Vonstad EK, Su X, Vereijken B, Solbjør M, Skjæret-Maroni N. Experiences of stroke survivors and clinicians with a fully immersive virtual reality treadmill exergame for stroke rehabilitation: a qualitative pilot study. Front Aging Neurosci 2021;13:735251 [FREE Full text] [doi: 10.3389/fnagi.2021.735251] [Medline: 34795576]
19. DiCarlo JA, Gheihman G, Lin DJ. Reimagining stroke rehabilitation and recovery across the care continuum: results from a design-thinking workshop to identify challenges and propose solutions. Archives of Physical Medicine and Rehabilitation 2021;102(8):1645-1657. [doi: 10.1016/j.apmr.2021.01.074] [Medline: 33556351]
20. Erikson A, Ranner M, Guidetti S, von Koch L. In search of self after stroke: a longitudinal qualitative study in the context of client-centred rehabilitation. Int J Qual Stud Health Well-being 2023;18(1):2282513 [FREE Full text] [doi: 10.1080/17482631.2023.2282513] [Medline: 38010156]
21. Bucki B, Spitz E, Baumann M. Emotional and social repercussions of stroke on patient-family caregiver dyads: analysis of diverging attitudes and profiles of the differing dyads. PLoS One 2019;14(4):e0215425 [FREE Full text] [doi: 10.1371/journal.pone.0215425] [Medline: 31013289]
22. Maggio M, Latella D, Maresca G, Sciarrone F, Manuli A, Naro A, et al. Virtual reality and cognitive rehabilitation in people with stroke: an overview. J Neurosci Nurs 2019;51(2):101-105. [doi: 10.1097/JNN.0000000000000423] [Medline: 30649091]
23. Johansen T, Matre M, Tornås S, Løvstad M, Ponsford JL, Olsen A, et al. I thought it would be difficult, but this is actually something I can do - experiences with virtual reality-based cognitive training in persons with TBI. Ann Med 2025;57(1):2490218 [FREE Full text] [doi: 10.1080/07853890.2025.2490218] [Medline: 40219760]
24. Bally E, Cheng D, van Grieken A, Ferri Sanz M, Zanutto O, Carroll A, et al. Patients' perspectives regarding digital health technology to support self-management and improve integrated stroke care: qualitative interview study. J Med Internet Res 2023;25:e42556 [FREE Full text] [doi: 10.2196/42556] [Medline: 37014677]
25. Brassel S, Power E, Campbell A, Brunner M, Togher L. Recommendations for the design and implementation of virtual reality for acquired brain injury rehabilitation: systematic review. J Med Internet Res 2021;23(7):e26344 [FREE Full text] [doi: 10.2196/26344] [Medline: 34328434]
26. Salim H, Lee PY, Sharif-Ghazali S, Cheong AT, Wong J, Young I, RESPIRE Collaboration. Developing an asthma self-management intervention through a web-based design workshop for people with limited health literacy: user-centered design approach. J Med Internet Res 2021;23(9):e26434 [FREE Full text] [doi: 10.2196/26434] [Medline: 34499039]
27. Yin R. Case Study Research: Design and Methods. Thousand Oaks, California: Sage; 2009.
28. Ridder H. The theory contribution of case study research designs. Bus Res 2017;10(2):281-305. [doi: 10.1007/s40685-017-0045-z]
29. Pereira JF, Santos J, Bertanha M. Rapid prototyping and health: an ergonomic footrest to improve the venous return. Product 2020;18(2):196-203. [doi: 10.4322/pmd.2020.031]

30. Bernhardt J, Hayward KS, Kwakkel G, Ward NS, Wolf SL, Borschmann K, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the stroke recovery and rehabilitation roundtable taskforce. Int J Stroke 2017;12(5):444-450 [FREE Full text] [doi: 10.1177/1747493017711816] [Medline: 28697708]
31. 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]
32. Platz T, Pinkowski C, van Wijck F, Kim I, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer test, action research arm test and box and block test: a multicentre study. Clin Rehabil 2005;19(4):404-411. [doi: 10.1191/0269215505cr832oa] [Medline: 15929509]
33. Persson HC, Alt Murphy M, Danielsson A, Lundgren-Nilsson, Sunnerhagen KS. A cohort study investigating a simple, early assessment to predict upper extremity function after stroke - a part of the SALGOT study. BMC Neurol 2015;15:92 [FREE Full text] [doi: 10.1186/s12883-015-0349-6] [Medline: 26084397]
34. Steen M, Manschot M, De Koning N. Benefits of co-design in service design projects. International Journal of Design 2011;5(2):53-60 [FREE Full text]
35. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
36. Younas A, Masih Y, Sundus A. Alternatives to ‘saturation’ for greater transparency in reporting of sample size decision-making in qualitative research. Evid Based Nurs 2025;28(2):77-80. [doi: 10.1136/ebnurs-2025-104292]
37. Creswell J. Qualitative Inquiry and Research Design: Choosing Among Five Traditions. Thousand Oaks: SAGE Publications; 1998.
38. Meydan C. The role of triangulation in qualitative research: converging perspectives. In: Principles of Conducting Qualitative Research in Multicultural Settings. Hershey, Pennsylvania: IGI Global; 2024:98-129.
39. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
40. 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 and Psychother Res 2020;21(1):37-47. [doi: 10.1002/capr.12360]
41. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
42. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 2008;51(1):S225-S239. [doi: 10.1044/1092-4388(2008/018)] [Medline: 18230848]
43. Xu S, Xu Y, Wen R, Wang J, Qiu Y, Chan CC. Virtual reality enhanced exercise training in upper limb function of patients with stroke: meta-analytic study. J Med Internet Res 2025;27:e66802 [FREE Full text] [doi: 10.2196/66802] [Medline: 39969977]
44. Maheta B, Kraft A, Interrante N, Fereydooni S, Bailenson J, Beams B, Keny, et al. Using virtual reality to improve outcomes related to quality of life among older adults with serious illnesses: systematic review of randomized controlled trials. J Med Internet Res 2025;27:e54452. [doi: 10.2196/54452] [Medline: 40009834]
45. Jayasree-Krishnan V, Ghosh S, Palumbo A, Kapila V, Raghavan P. Developing a framework for designing and deploying technology-assisted rehabilitation after stroke: a qualitative study. Am J Phys Med Rehabil 2021;100(8):774-779. [doi: 10.1097/PHM.0000000000001634] [Medline: 33141773]
46. National Clinical Guideline Centre (UK). Planning and delivering stroke rehabilitation. In: Stroke Rehabilitation: Long Term Rehabilitation After Stroke [internet]. UK: Royal College of Physicians (UK); 2013.
47. Gustavsson M, Kjörk EK, Erhardsson M, Alt Murphy M. Virtual reality gaming in rehabilitation after stroke - user experiences and perceptions. Disabil Rehabil 2022;44(22):6759-6765 [FREE Full text] [doi: 10.1080/09638288.2021.1972351] [Medline: 34465269]
48. Pereira F, Bermúdez I Badia S, Jorge C, Cameirão MS. The use of game modes to promote engagement and social involvement in multi-user serious games: a within-person randomized trial with stroke survivors. J Neuroeng Rehabil 2021;18(1):62 [FREE Full text] [doi: 10.1186/s12984-021-00853-z] [Medline: 33853609]
49. Oppici L, Bērziņa G, Hestetun-Mandrup AM, Løvstad M, Opheim A, Pacheco MM, PEER-HOMEcare consortium, et al. A scoping review of preclinical environmental enrichment protocols in models of poststroke to set the foundations for translating the paradigm to clinical settings. Transl Stroke Res 2025;16(5):1850-1873. [doi: 10.1007/s12975-025-01335-3] [Medline: 39913056]
50. Khan A, Imam YZ, Muneer M, Al Jerdi S, Gill SK. Virtual reality in stroke recovery: a meta-review of systematic reviews. Bioelectron Med 2024;10(1):23. [doi: 10.1186/s42234-024-00150-9]
51. Levin MF. Can virtual reality offer enriched environments for rehabilitation? Expert Rev Neurother 2011;11(2):153-155. [doi: 10.1586/ern.10.201] [Medline: 21306202]
52. Sokołowska B. Being in virtual reality and its influence on brain health-an overview of benefits, limitations and prospects. Brain Sci 2024;14(1):72 [FREE Full text] [doi: 10.3390/brainsci14010072] [Medline: 38248287]

53. Ceradini M, Losanno E, Micera S, Bandini A, Orlandi S. Immersive VR for upper-extremity rehabilitation in patients with neurological disorders: a scoping review. J Neuroeng Rehabil 2024;21(1):75 [FREE Full text] [doi: 10.1186/s12984-024-01367-0] [Medline: 38734690]
54. Hestetun-Mandrup AM, Hamre C, Lund A, Martinsen ACT, He H, Pikkarainen M. Exploring people with stroke's perceptions of digital technologies in post-stroke rehabilitation - a qualitative study. Disabil Rehabil 2025:1-10 [FREE Full text] [doi: 10.1080/09638288.2025.2504615] [Medline: 40382688]
55. Hestetun-Mandrup AM, Toh ZA, Oh HX, He H, Martinsen ACT, Pikkarainen M. Effectiveness of digital home rehabilitation and supervision for stroke survivors: a systematic review and meta-analysis. Digit Health 2024;10:20552076241256861 [FREE Full text] [doi: 10.1177/20552076241256861] [Medline: 38832099]
56. Olana DD, Abessa TG, Lamba D, Triccas LT, Bonnechere B. Effect of virtual reality-based upper limb training on activity of daily living and quality of life among stroke survivors: a systematic review and meta-analysis. J Neuroeng Rehabil 2025;22(1):92 [FREE Full text] [doi: 10.1186/s12984-025-01603-1] [Medline: 40269877]
57. Rodríguez-Hernández M, Polonio-López B, Corregidor-Sánchez AI, Martín-Conty JL, Mohedano-Moriano A, Criado-Álvarez JJ. Can specific virtual reality combined with conventional rehabilitation improve poststroke hand motor function? A randomized clinical trial. J Neuroeng Rehabil 2023;20(1):38 [FREE Full text] [doi: 10.1186/s12984-023-01170-3] [Medline: 37016408]
58. Dixit, Phalswal U, Kalal N, Srivastava SP. Effectiveness of virtual reality-supported exercise therapy in improving upper extremity function and activities of daily living among patients after stroke: a systematic review of randomized control trials. Osong Public Health Res Perspect 2024;15(3):189-200 [FREE Full text] [doi: 10.24171/j.phrp.2023.0148] [Medline: 38988022]
59. Soleimani M, Ghazisaeedi M, Heydari S. The efficacy of virtual reality for upper limb rehabilitation in stroke patients: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2024 May 24;24(1):135 [FREE Full text] [doi: 10.1186/s12911-024-02534-y] [Medline: 38790042]
60. Rodrigues P, Quaresma C, Costa M, Luz F, Fonseca MM. Virtual reality-based telerehabilitation for upper limb recovery post-stroke: a systematic review of design principles, monitoring, safety, and engagement strategies. arXiv:2501.06899 2025:00 [FREE Full text] [doi: 10.48550/arXiv.2501.06899]
61. Gkintoni E, Vassilopoulos SP, Nikolaou G. Next-generation cognitive-behavioral therapy for depression: integrating digital tools, teletherapy, and personalization for enhanced mental health outcomes. Medicina (Kaunas) 2025;61(3):431 [FREE Full text] [doi: 10.3390/medicina61030431] [Medline: 40142242]
62. Pourliaka T, Tsiakiri A, Proios H. Exploring discharge preparedness of patients with stroke: a qualitative descriptive study. Creat Nurs 2025;31(4):440-450. [doi: 10.1177/10784535251320005] [Medline: 40017407]
63. Martínez-Cid S, Vallejo D, Herrera V, Schez-Sobrino S, Castro-Schez JJ, Albusac JA. Explainable AI-driven decision support system for personalizing rehabilitation routines in stroke recovery. Prog Artif Intell 2025:1-23. [doi: 10.1007/s13748-024-00357-6]
64. McDicken JA, Elliott E, Blayney G, Makin S, Ali M, Larner AJ, VISTA-Cognition Collaborators. Accuracy of the short-form Montreal Cognitive Assessment: systematic review and validation. Int J Geriatr Psychiatry 2019 Oct;34(10):1515-1525 [FREE Full text] [doi: 10.1002/gps.5162] [Medline: 31243810]

Abbreviations

AI: artificial intelligence
MoCA: Montreal Cognitive Assessment
TA: thematic analysis
VR: virtual reality

Edited by A Kushniruk; submitted 10.Jun.2025; peer-reviewed by D Hobbs, A Glaessel; comments to author 06.Aug.2025; revised version received 28.Aug.2025; accepted 25.Nov.2025; published 15.Jan.2026.

Please cite as:
Saidi T, Hestetun-Mandrup AM, Pavel N, Rosseland IKH, Selmer-Olsen KG, Mørk NS, Bergheim Å, Pikkarainen MA
Challenges and Benefits of Virtual Reality in Home-Based Poststroke Rehabilitation: Co-Design Qualitative Study
JMIR Hum Factors 2026;13:e78798
URL: https://humanfactors.jmir.org/2026/1/e78798
doi:10.2196/78798
PMID:

Privacy Fact Sheets for Mitigating Disease-Related Privacy Concerns and Facilitating Equal Access to the Electronic Health Record: Randomized Controlled Trial

Niklas von Kalckreuth, MSc; Markus A Feufel, MSc, PhD

Division of Ergonomics, Department of Psychology and Ergonomics, Technische Universität Berlin, Straße des 17. Juni 135, 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, Germany

Abstract

Background: The German electronic health record (EHR) aims to enhance patient care and reduce costs, but users often worry about data privacy and security. To mitigate disease-related privacy concerns, for instance, surrounding stigmatized diseases, we test the effect of privacy fact sheets (PFSs)—a concise but comprehensive transparency feature designed to increase users' perceived control over their data—on increasing EHR use in a simulated online study.

Objective: The study aimed to investigate whether displaying a PFS shortly before upload decisions must be made mitigates disease-related privacy concerns and makes uploads more likely.

Methods: In an online survey study, 393 German participants from the recruitment platform Prolific were asked to interact with a randomly assigned medical report that varied systematically in terms of disease-related stigma (high vs low) and time course (TC; acute vs chronic). They were then asked to decide whether to upload the report to an EHR click dummy, while we systematically varied the presentation of privacy information (PFS vs no PFS). Participants were randomly (single-blinded) assigned to one of the 2×2×2 conditions (stigma, TC, privacy information).

Results: All 393 participants were randomly assigned to one of the following groups: low, acute, no PFS (n=52, 13.2%); low, chronic, no PFS (n=45, 11.5%); high, acute, no PFS (n=46, 11.7%); high, chronic, no PFS (n=55, 14%); low, acute, PFS (n=44, 11.2%); low, chronic, PFS (n=41, 10.4%); high, acute, PFS (n=56, 14.2%); and high, chronic, PFS (n=54, 13.7%). The results show that, in general, upload behavior is negatively influenced by disease-related stigma (odds ratio [OR] 0.130; *P*<.001) and positively influenced when a PFS is given (OR 4.527; *P*<.001). This increase was particularly pronounced for stigmatized diseases (OR 5.952; *P*=.006), but independent of the TC of the diseases.

Conclusions: Our results demonstrate that PFSs may help to increase EHR uploads in people interacting with a realistic EHR click dummy, by mitigating privacy concerns in scenarios involving stigmatized diseases. Results further indicate that a PFS is mainly relevant and effective for people with increased privacy risk perceptions, whereas they neither benefit nor hurt others. Thus, implementing PFSs may increase the likelihood that users who perceive high privacy risks when confronted with sensitive or stigmatized health information decide to upload their data to the EHR, ultimately increasing digital health equity.

Trial Registration: Deutsches Register Klinischer Studien DRKS00033652; https://drks.de/search/de/trial/DRKS00033652

(JMIR Hum Factors 2026;13:e71124) doi:10.2196/71124

KEYWORDS

adoption; disease characteristics; EHR; electronic health record; mHealth; mobile health; stigma; time course; transparency; transparency feature; privacy concern

Introduction

Background

The electronic health record (EHR) is one key element in the digital transformation of health care systems because it allows patients' health data (eg, diagnoses, therapies, vaccinations, and medication plans) to be readily documented, exchanged, and viewed by various stakeholders [1-6]. The resulting network of care providers can make patient treatment more effective, safer, and faster across institutions [7]. For instance, preexisting conditions, intolerances, and medication plans can be taken into account during diagnosis and treatment to prevent adverse medication interactions, duplicate diagnoses, overtreatment, and undertreatment [5]. Also, it is hoped that physicians will spend less time on obtaining patients' medical history thanks to the EHR, which they could devote to actual patient treatment [8]. In Germany, approximately 90% of residents are covered by statutory health insurance and thus received an EHR account by default in January 2025, unless they opted out [9-11]. From October 1, 2025, health care providers are required to use the EHR infrastructure for documentation and data exchange [11]. A prerequisite for the potential of EHRs to realize is user engagement. Specifically, the Patient Data Protection Act mandates that patients maintain sole control of their data, allowing them to decide which information is stored in the EHR, who has access to it, and which data are to be deleted [5]. Consequently, the success of the EHR hinges on patients' actual use of the technology. But national and international studies indicate that many patients remain skeptical toward the EHR, mainly due to concerns arising from limited trust in both data security (ie, the technical and organizational measures that protect personal data) and data privacy (ie, the rights and rules governing how personal data is collected, used, and shared) [8,12-15]. In Germany, these concerns were intensified due to reported security issues, which were addressed and resolved during the rollout [16,17]. However, concerns about EHR use are not static in nature. Our previous studies suggest that participants adapt their attitudes toward EHR dynamically. For instance, we could show that the perceived risks and benefits of using EHR are related to disease-specific privacy concerns, such as the stigma and the time course (TC) of diseases, that is, the more permanent and risky nature of data associated with chronic conditions [18-20]. Previous studies have also shown that existing communication strategies, for instance, by health insurers or the Federal Ministry of Health, have not been sufficient to effectively convey the core functions and data protection mechanisms of the German EHR [13,14,21]. Against this background, there is a need for concise and comprehensible communication strategies that can transparently explain data privacy and data security measures, thereby strengthening user trust in the EHR. In this study, we investigate the effect of a privacy fact sheet (PFS)—a concise but comprehensive transparency feature—on increasing EHR use and, specifically, to what extent the PFS can mitigate disease-related privacy concerns and increase the upload of medical reports to the EHR.

Prior Work

"Notice and choice" is the most widely used framework for ensuring data privacy worldwide [22,23]. As its name suggests, it consists of 2 components: privacy notices and privacy choices. Whereas *privacy notices* explain how personal data are collected, processed, and shared with third parties, *privacy choices* give users control over various aspects of these practices, including the decision to start and terminate them [22]. Various studies indicate that, if informed by privacy notices, users are empowered to choose IT systems that match their preferences, typically those with high data security and privacy standards, and avoid less secure ones [24,25]. But the current formats used for privacy notices, most commonly *privacy policies*, tend to provide rather detailed information and often use legal jargon [26-30], which aims to maximize legal protection of IT providers rather than to transparently inform users [31]. Research has shown that overly lengthy and complex privacy policies may ultimately serve as a "red flag," leading users to lose trust in the provider, if not to discontinue technology use altogether [32,33]. Consequently, concise, easy-to-understand privacy notices are a prerequisite for users to adopt digital health technologies such as the EHR [33-35].

In contrast to full-blown privacy policies, the shorter *transparency features* have been shown to be an effective type of privacy notice, because they provide a brief and easy-to-understand overview of data privacy and data security measures and are meant to inform rather than to provide legal assurance [36,37]. Recent work also shows that such transparency features can decrease privacy concerns and increase EHR acceptance by strengthening users' perceived control over personal data—a construct that describes the extent to which users feel able to understand, oversee, and control how their personal data are handled and that plays a central role in privacy calculus models [7,38-40]. Stronger perceptions of control are known to reduce privacy concerns and foster trust in digital health and online technologies [7,38,39,41].

Empirical studies across domains suggest that transparency features may influence user behavior. For instance, a study in the eCommerce domain demonstrated that displaying a transparency feature positively influences purchase numbers [37]. But increased use does not (only) depend on the format of the privacy notice; it is also influenced by the contents it provides, including the efficacy of the mentioned data protection measures and privacy choices, and its timing, that is, when the privacy notice is given to users [42]. At the same time, studies in other digital contexts have reported mixed results regarding the behavioral impact of privacy notices and transparency features, suggesting that their effectiveness may depend on contextual factors such as perceived privacy risk or data sensitivity [36,43]. Our previous studies have shown that a transparency feature with a concise but comprehensive summary of all relevant contents—which we refer to as a PFS—positively influences EHR use when given shortly before the upload process [40,44]. In addition, we could show that a patient-centered framing of these contents that specifies what users can do to control the EHR and emphasizes their perceived control over personal data (eg, you can control all of your data) has the biggest effect on EHR adoption [44].

As stated at the outset, in another line of studies, we have shown that privacy concerns, intention to use the EHR, and upload behavior are influenced by the characteristics of diseases, in particular by disease-related stigma and TC [18-20]. Disease-specific stigma has been shown to have an inhibiting influence on upload behavior and to increase the risks people perceive when they are asked to upload information related to these diseases to the EHR [18,19]. Conversely, the TC of diseases (ie, whether diseases are chronic rather than acute) tends to increase both privacy concerns and intention to use the EHR. That is, patients with chronic conditions recognize a greater value in using the EHR but have heightened privacy concerns when it comes to uploading chronic conditions to the EHR [20]. In this study, we aim to merge these two lines of studies to validate and extend the positive effect of a patient-framed PFS on users' decision to upload diseases to the EHR click dummy when disease-specific privacy concerns are systematically varied.

Aim of This Research and Approach

In this study, we test whether displaying a patient-framed PFS, developed in our previous studies [40,44], shortly before the decision to upload a medical report must be made increases the likelihood that users upload medical reports to the EHR for diseases that vary along 2 dimensions: TC and disease-specific stigma. After describing the methods and results, we discuss the implications, reflect on the study's limitations, and conclude with a reflection on the objective of this study.

Methods

Ethical Considerations

This study was approved by the Ethics Committee of the Department of Psychology and Ergonomics at Technische Universität Berlin (tracking number: AWB_KAL_1_230206_Erweiterungsantrag). The study is registered as a randomized controlled trial at Deutsches Register Klinischer Studien (DRKS00033652). Participants volunteered to participate in the survey, and written informed consent was required to participate. On the first page of the survey, participants were told about the experimenter, the study purpose, what data were to be collected during the study, and where and for how long they would be stored. Also, participants had the possibility to download a PDF with the study information. Hence, participants were informed about the duration of the survey (approximately 8 min) as well as the compensation for participation. All data were collected and stored in an anonymized form. No directly identifying personal information was collected. Data were processed confidentially and used exclusively for research purposes.

Participants

The online study was conducted between April 15, 2024, and May 16, 2024. Based on an a priori power analysis for a logistic regression using G*Power (version 3.1.9.7) with disease-related stigma (high vs low), TC (acute vs chronic), and privacy information (PFS vs no PFS) as binomial distributed predictors, a false positive rate α of .05, a power β of .80, an estimated odds ratio (OR) for the predictor with the smallest expected effect size (TC) of 1.7 (derived from the prestudy with n=80 participants), and a probability of the outcome (upload decision) under the null hypothesis of 0.5, reflecting a conservative assumption, we aimed for a sample size of 363 participants. To ensure this target was met, we oversampled participants by 30%, resulting in a total sample of 471 individuals. Oversampling accounted for potential exclusions due to failed attention checks, study dropouts, self-reported invalid data (approximately 20%), as indicated in preliminary studies [40,44], and prior medical histories with the diseases used in the study (approximately 10%, based on prior findings) [18]. Individuals 18 years and older residing in Germany were allowed to participate in the study, as the content and questions of the study were designed to fit the context of the German EHR. Another prerequisite was that participants had no personal previous experience (own illness) with the diseases mentioned in the medical reports we used for this study, as the handling of stigmatized diseases by affected persons is different from that of unaffected persons [45]. Sampling was conducted through Prolific, a crowdsourcing platform used to recruit participants for online surveys and experiments, known for its diverse participant pool and high data quality [46]. Participation was compensated with 1.78€ (US $2.09) for 8 minutes, which corresponds to the German minimum wage. The mean value of the processing time was 8:47 minutes (SD 3:57 min), and the median was 8 minutes. A total of 471 individuals participated in the study.

Design

The experimental design replicates the disease-related manipulations from our previous study [18] and extends them by testing whether transparent privacy information can reduce disease-related privacy concerns and influence upload behavior. This approach also extends prior privacy calculus–based work by focusing on actual behavior rather than intentions, as behavioral measures provide stronger ecological validity and better capture situational influences on privacy decisions [47,48]. We used a 2×2×2 between-subject study design with the 3 independent variables (IVs): stigmatization potential (SP), TC, and privacy information. Each participant was assigned to one unique combination of these conditions. As in preliminary studies, SP (high vs low) and TC (acute vs chronic) were manipulated by displaying the diagnoses of a disease with the respective characteristics [18,19]. Additionally, privacy information (PFS vs no PFS) was manipulated by either displaying a PFS during the upload process or not. In preliminary studies, participants associate disease-related stigma with high risks [18,19], and consequences could arise in areas related to personal lifestyle, occupation, and social life if medical findings became known [19,45,49,50]. Furthermore, previous studies show that participants perceive the upload of diseases with a chronic TC as more beneficial than the upload of acute diseases [8,19,20]. Participants were randomly (single-blinded) assigned to one of the conditions in parallel (simple randomization, ratio: 1:1:1:1:1:1:1:1) using LimeSurvey's built-in "rand" function. The randomization process was fully automated within LimeSurvey, ensuring allocation concealment throughout data collection. Neither the participants nor the researchers conducting the data collection were aware of the assigned condition at any point prior to or during data collection. The dependent variable was the decision to upload the medical report, that is, whether participants were willing to upload the medical findings to the EHR [18,19,40]. Reporting of this randomized controlled trial followed the CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines; a completed CONSORT-eHEALTH checklist is provided in Checklist 1.

Materials

Following a common practice in technology acceptance studies [51,52], we used a case vignette to represent a typical situation in which an EHR app may be used. In particular, the case vignette depicted a situation where the participant has recently started using an EHR app and is now faced with the decision to upload a medical finding to their EHR (see Multimedia Appendix 1). Additionally, the disease/injury was described in lay terms with 1 to 3 sentences (see Multimedia Appendix 2). The stimuli used in the study were realistic but specially created for the purpose of the study. The medical reports were provided by 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. This was done with the permission of the institutions concerned. In selecting the diseases, both the related stigma and their TC were systematically varied. Disease-related stigma covered different risks for professional and social life, such as tests for sexually transmitted diseases (ie, gonorrhea and HIV) [53-56] and fractures or rheumatoid arthritis as diseases with low stigma. To reflect different TCs, diseases were divided according to an acute TC (eg, wrist fracture and gonorrhea) and a chronic one (eg, rheumatoid arthritis and HIV). Furthermore, diseases were selected to occur regardless of age, meaning they can affect individuals across different age groups, 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 perceived stigma and TC.

Table . Diseases used as stimuli, categorized by SP[a] and TC[b].

SP and TC	Acute	Chronic
Low	Fractured wrist	Rheumatoid arthritis
High	Gonorrhea	HIV infection

[a] SP: stigmatization potential.

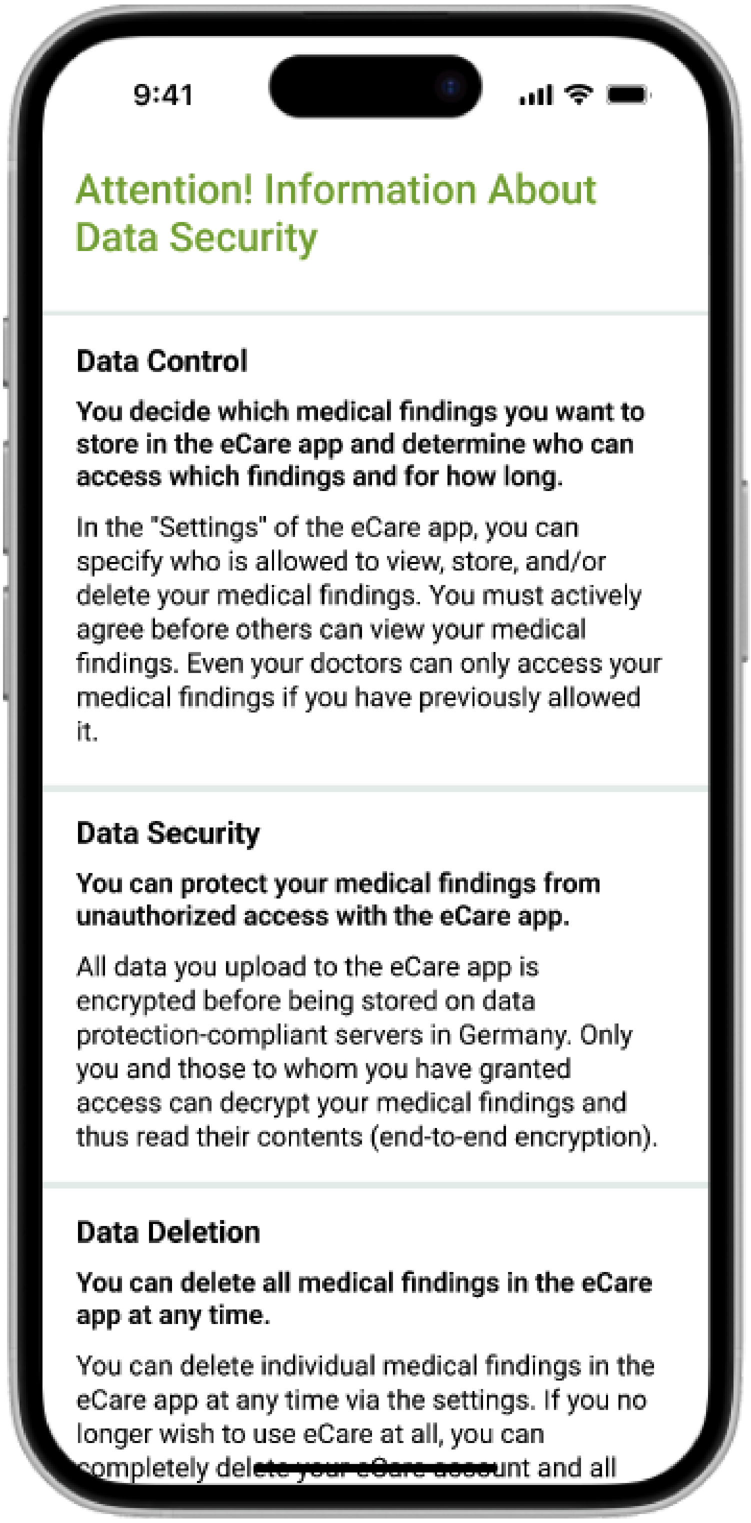
[b] TC: time course.

As in a previous study [19], an interactive prototype (a so-called click dummy) was used, which we created after the mobile EHR app of a German health insurance company (the BARMER) using software for interface design (FIGMA). This prototype allows for a realistic interaction with an EHR. Specifically, the prototype gave participants the ability to upload findings, grant or revoke permissions to view findings, and create medication plans. Only the “Upload findings” function was used in this study.

We used the most effective PFS that we identified based on preliminary studies [40,44], which was marked by a concise but comprehensive content and a patient-centered framing, that is, a description of what the EHR allows its users to do to control their data (eg, you can control all of your data) rather than what it does for them (eg, the EHR keeps all of your data safe). Figure 1 shows the English version of the PFS used. The English translation of the full text can be found in the Multimedia Appendix 3.

We used LimeSurvey (version 3.28.66+230719) to create and conduct a 9-page online survey. The EHR prototype was embedded 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 previous studies, we tested the effect of the IVs by querying the perceived risk and perceived benefit of uploading findings to the EHR using validated items [18-20]. Also, we assumed that people perceived more risk when the SP was high and more benefit when the TC was chronic [19,20]. Perceived risk and perceived benefit were measured using a 7-point Likert scale ranging from 1 (“Strongly disagree”) to 7 (“Strongly agree”). The decision to upload the finding was measured using a validated dichotomous item (yes/no) [18,19].

Figure 1. Privacy fact sheets used in the study.



Attention! Information About Data Security

Data Control

You decide which medical findings you want to store in the eCare app and determine who can access which findings and for how long.

In the "Settings" of the eCare app, you can specify who is allowed to view, store, and/or delete your medical findings. You must actively agree before others can view your medical findings. Even your doctors can only access your medical findings if you have previously allowed it.

Data Security

You can protect your medical findings from unauthorized access with the eCare app.

All data you upload to the eCare app is encrypted before being stored on data protection-compliant servers in Germany. Only you and those to whom you have granted access can decrypt your medical findings and thus read their contents (end-to-end encryption).

Data Deletion

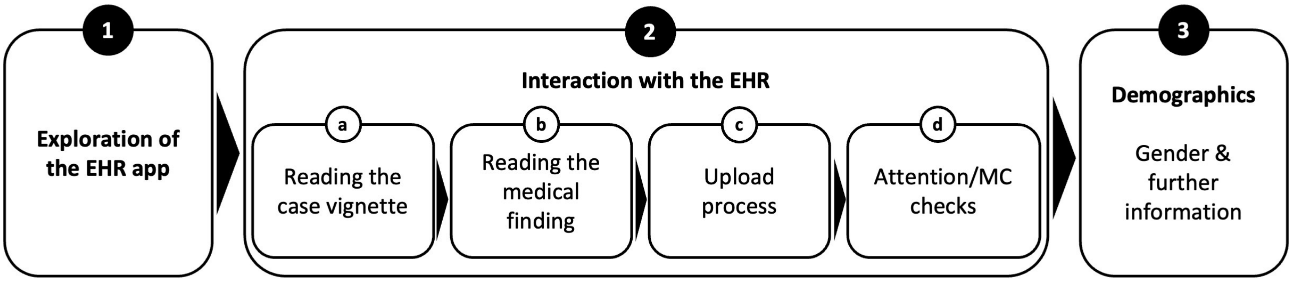
You can delete all medical findings in the eCare app at any time.

You can delete individual medical findings in the eCare app at any time via the settings. If you no longer wish to use eCare at all, you can completely del~~ete your eCare acco~~unt and all

Procedure

The study procedure is shown in Figure 2. The survey consisted of 3 parts. After giving their written informed consent according to the Declaration of Helsinki of the World Medical Association, (1) participants had several minutes to interact with the EHR prototype. (2a) Participants then read a randomly selected case vignette addressing the use of the EHR in the context of uploading a medical finding. (2b) Additionally, the participants read the medical findings of the respective illness (low or high SP and acute or chronic TC, depending on the experimental group), as well as a brief description of the respective disease. (2c) Afterwards, as part of the upload process, the participants were asked to select the medical finding for upload. Depending on the experimental group, either a PFS was displayed before the disease could be selected or not. Participants then decided whether they wanted to upload the report to their EHR. (2d) After uploading, participants who were shown a PFS were asked a question about the content of the texts to ensure that the texts were read (attention check), and all participants were asked about the perceived privacy risks and benefits of uploading the report (manipulation check). (3) The survey was completed with the collection of demographic characteristics (age, gender, education level, and experience with mobile health [mHealth] apps) as control variables, as well as the opportunity for participants to declare their responses invalid due to lack of care in processing them (see Multimedia Appendix 4 for the questionnaire).

Figure 2. Overview of the study design. EHR: electronic health record; MC: manipulation check.



Hypotheses

As mentioned above, we hypothesize that diseases with high stigma would result in a high perceived risk and a chronic TC in a high perceived benefit. Hence, we hypothesize that the upload decision is negatively influenced by high disease stigma (H1) and positively influenced by a chronic TC (H2). Based on previous studies, we also assume that a PFS will generally increase upload behavior compared to the no PFS condition (H3). More specifically, regarding the aim of our study, if a PFS can mitigate disease-related concerns, we hypothesize that showing a PFS mitigates the negative influence of disease-related stigma on the upload decision (H4). Furthermore, we hypothesize that the positive influence of a chronic TC on the upload decision will be enhanced by the presence of a PFS, as it enhances perceived benefits related to long-term health management (H5). Textbox 1 provides an overview of the hypotheses regarding the IVs.

Textbox 1. Overview of the hypotheses.

The following were the hypotheses regarding the independent variables:

- H1: The number of uploads of medical findings to the electronic health record (EHR) is lower for diseases that are stigmatized compared to those that are nonstigmatized.
- H2: The number of uploads of medical findings to the EHR is higher for chronic diseases compared to acute diseases.
- H3: The number of uploads of medical findings to the EHR is higher when transparency regarding data privacy and security is high compared to when it is low.
- H4: The increase in the number of uploads when showing a privacy fact sheet (PFS) is higher for stigmatized diseases than for nonstigmatized diseases.
- H5: The increase in the number of uploads when showing a PFS is higher for chronic diseases than for acute diseases.

Analyses

We cleaned and analyzed the data using RStudio (version 2023.09.1+494). The analysis regarding the manipulation checks of perceived privacy risks and benefits was performed using *t* tests, a statistical method used to compare the means of 2 groups. The influence of the IVs (disease-specific stigma, TC, displaying a PFS) and the interaction effects between stigma and the display of a PFS, as well as between TC and the display of a PFS on the upload decision, were tested using multiple logistic regression with dummy coding, a method used to model the probability of a binary outcome based on one or more predictor variables.

We also included a robustness check of the results regarding the upload decision. To control for potential influences of demographic and interindividual variables that could bias coefficients and *P* values, we used multiple logistic regression. To not bias *P* values as a result of controlling, we only included variables in the model that have been shown to have a causal relationship with the IVs (ie, causal confounders): age, education level, and experience with the technical system [38,57,58]. *P* values were adjusted for multiple testing using the Benjamini-Hochberg procedure [59].

Results

Survey Characteristics

A total of 471 observations were collected. A total of 78 (16.5%) records were excluded, of which 70 (14.9%) were excluded because of incomplete questionnaires, 4 (0.85%) because participants failed the attention check, and 4 (0.85%) because responses were marked as invalid by participants. A sample of 393 observations (156 female participants, 231 male participants, and 6 with no information) was used for further analysis. Figure 3 shows the participation and distribution process according to the guidelines of the CONSORT (Consolidated Standards of Reporting Trials) statement [60].

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) flow chart. PFS: privacy fact sheet; SP: stigma potential; TC: time course.

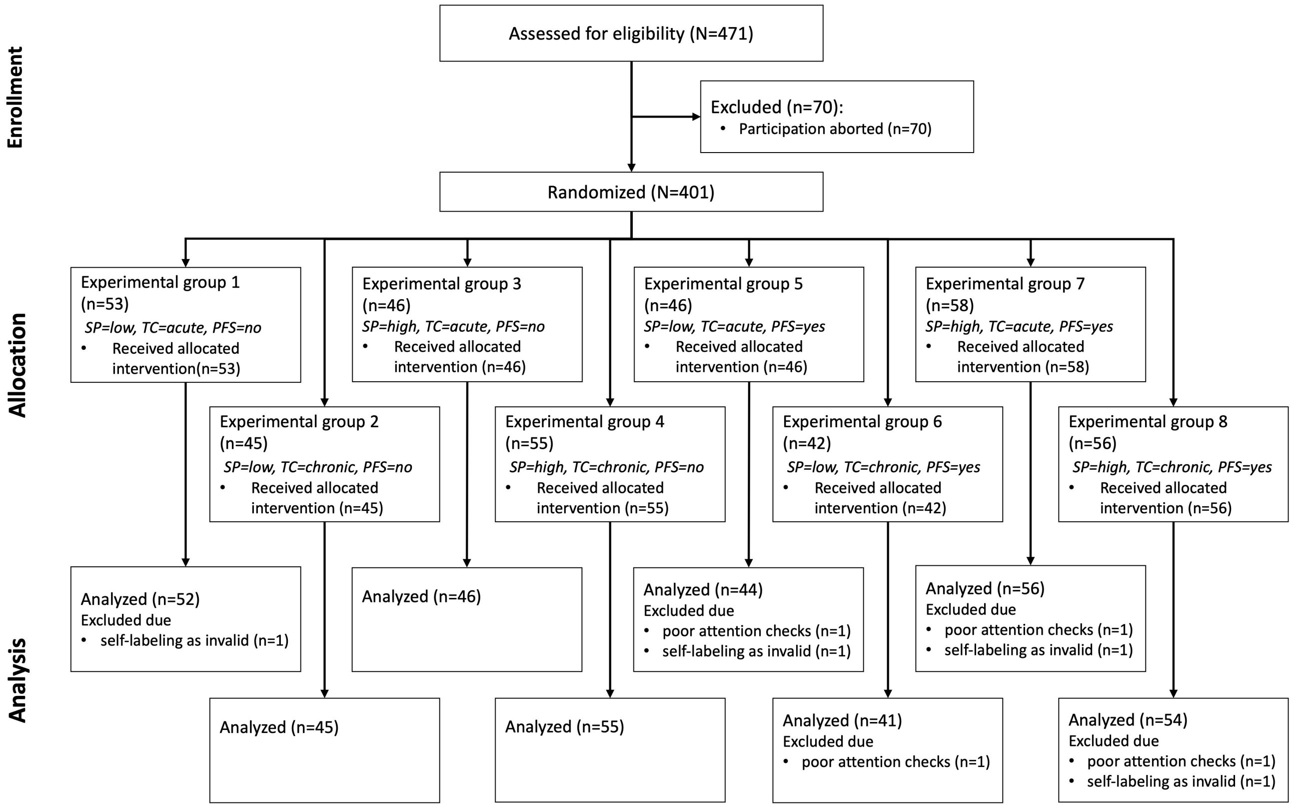


Table 2 summarizes the demographic characteristics of the entire sample. The demographic characteristics of the subsamples for each experimental group are shown in the Multimedia Appendix 5.

Table . Demographic data of the sample (n=393).

Demographic characteristic	Respondents
Age (y), mean (SD)	31.67 (9.94)
Sex, n (%)	
Female	156 (39.7)
Male	231 (58.8)
No answer	6 (1.5)
Education, n (%)	
No degree	11 (2.8)
High school/vocational education	179 (45.5)
Bachelor	102 (26.0)
Master	90 (22.9)
PhD	11 (2.8)
Experience with mHealth[a] apps, n (%)	
No use	226 (57.5)
Regular use	167 (42.5)

[a] mHealth: mobile health.

Risk and Benefit Perception

Similar to the preliminary studies [18,19], risk and benefit perception of uploading served as a manipulation check to test the validity of our manipulation (ie, the medical reports) with respect to the perception of risk (stigma) and benefit (TC). As expected, uploading medical findings of stigmatized diseases was perceived as riskier than those of nonstigmatized diseases (low: mean 3.88, SD 1.68; high: mean 5.15, SD 1.6; t_{391}=7.648; *P*<.001). Consequently, we assume that our risk manipulation was successful. There was no significant difference in the perceived benefit regarding the TC of the disease (acute: mean

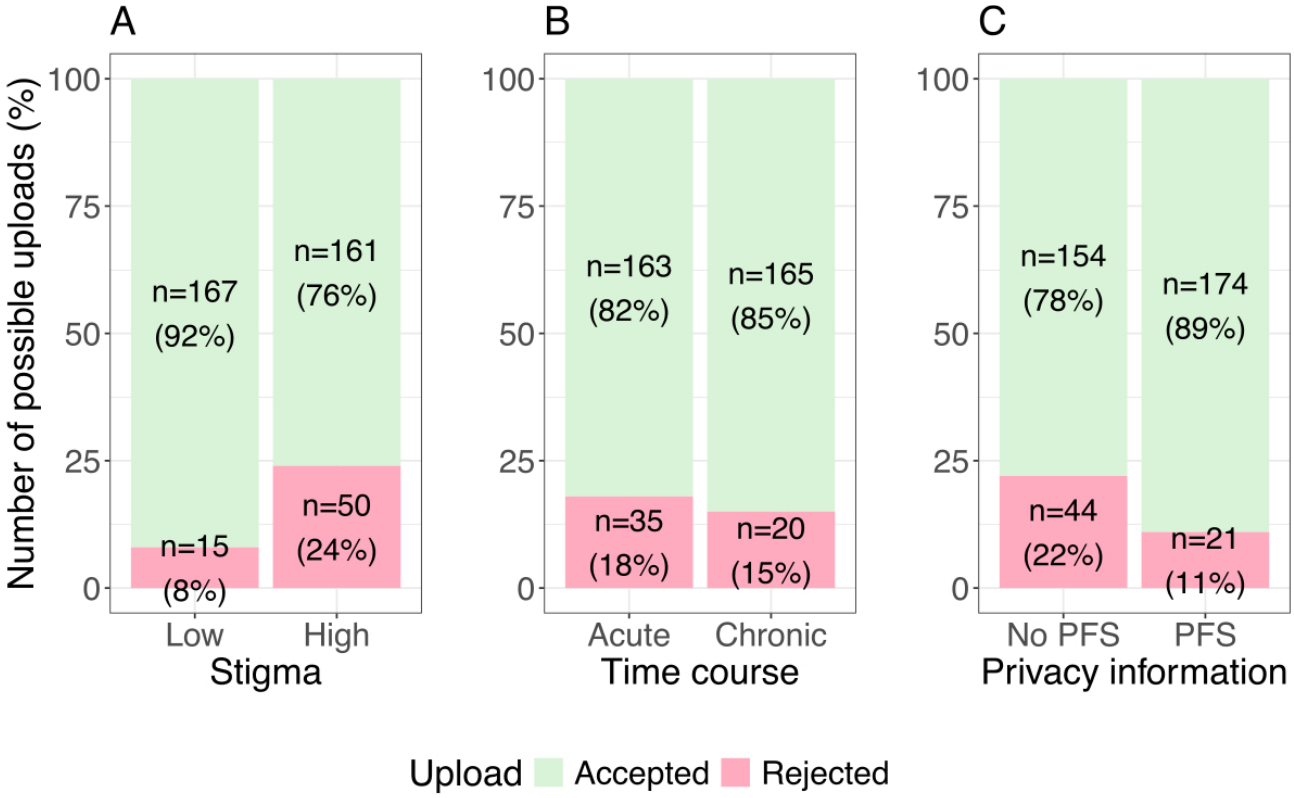
5.63, SD 1.41; chronic: mean 5.72, SD 1.16; t_{391}=0.703; *P*=.483). Consequently, we assume that our benefit manipulation was not successful.

Additionally, we analyzed the benefit perception in relation to stigma and the risk perception in relation to TC, even though these were not part of the initial manipulation checks. Uploading medical findings of nonstigmatized diseases was perceived as more beneficial than those of stigmatized diseases (low: mean 5.84, SD 1.16; high: mean 5.54, SD 1.39; t_{391}=2.345; *P*=.02). Furthermore, uploading medical findings of chronic diseases into the EHR was perceived as riskier than those of acute diseases (acute: mean 4.31, SD 1.83; chronic: mean 4.82, SD 1.64; t_{391}=2.893; *P*=.004).

Upload Behavior

Upload behavior was negatively associated with disease-related stigma (*z*=4.568; *P*<.001), thus supporting H1. Specifically, when stigma was high, it was more than seven times less likely that the report was uploaded to the EHR (76.3%, 161/211) than when stigma was low (91.8%, 167/182). TC of the disease was not associated with the decision to upload a report (z=0.877; *P*=.38). Consequently, H2 is rejected. The PFS was positively associated with the decision to upload a medical report to the EHR (*z*=3.298; *P*<.001), supporting H3. When a PFS was given, participants were more than 4 times as likely to upload the diagnosis to their EHR (89.2%, 174/195) than when a PFS was not given (77.7%, 154/198). The absolute number of uploads is shown in Figure 4 as a function of the IVs disease-related stigma, TC, and PFS vs no PFS.

Figure 4. Number of uploads to the electronic health record as a function of disease-related (A) stigma, (B) time course, and (C) privacy information. PFS: privacy fact sheet.



We also tested for interaction effects between stigma and privacy information, as well as between TC and privacy information, to explore potential moderating effects. The interaction between stigma and privacy information was significant (*z*=2.734; *P*=.006), indicating that the increase in the number of uploads when showing a PFS is higher for stigmatized diseases than for nonstigmatized diseases, thus supporting H4 (see Figure 5A).

In contrast, the interaction between TC and privacy information was not significant (*z*=0.094; *P*=.92), suggesting that displaying a PFS did not differentially impact the upload decision based on whether the disease was acute or chronic (see Figure 5B). Consequently, H5 is rejected. The summary of the results of the logistic regression is shown in Table 3.

Figure 5. Number of uploads to the electronic health record as interaction between stigma and privacy information (A) and time course and privacy information (B). PFS: privacy fact sheet.

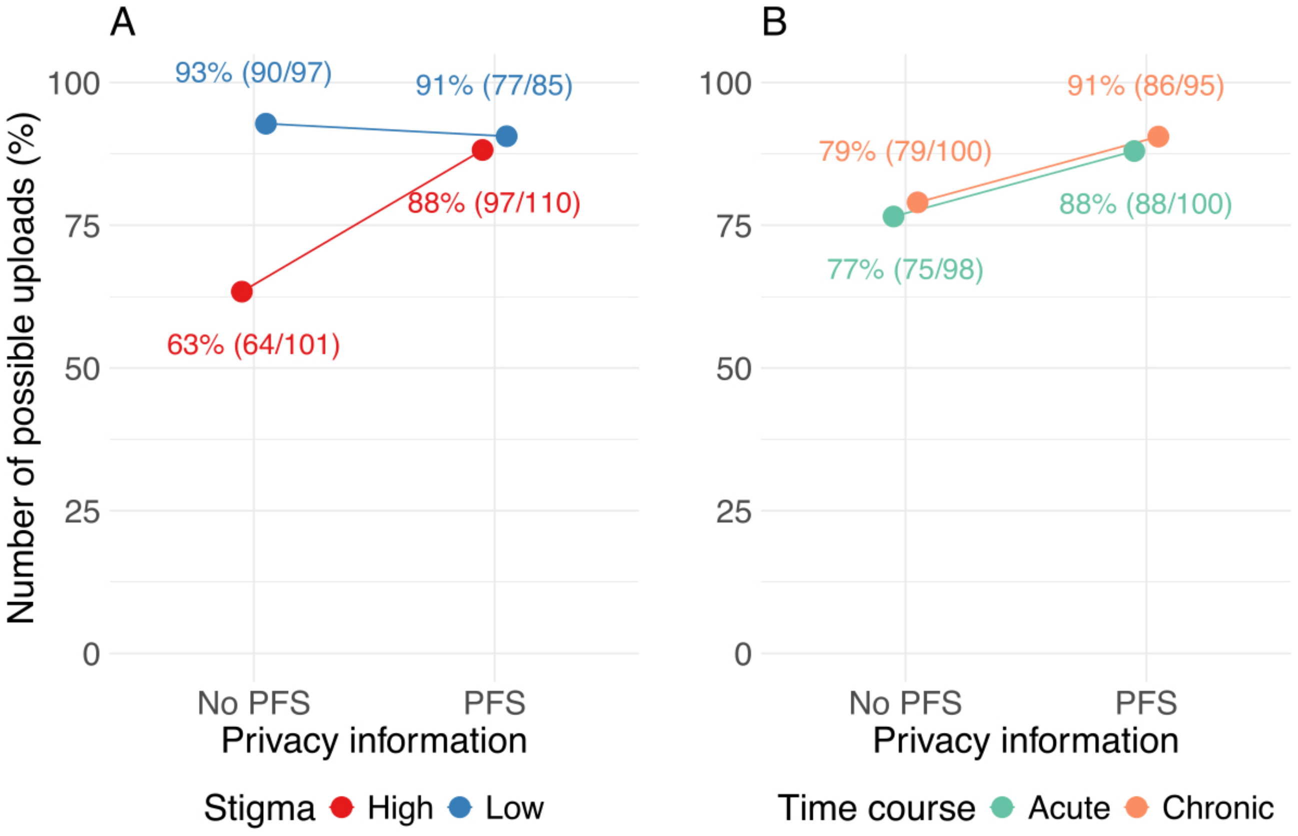


Table . Results of the logistic regression.[a]

Variable	z value	P value	ORs[b] and 95% CIs		
			Lower	OR	Upper
Stigma (high vs low)	4.568	<.001	0.050	0.130	0.296
TC[c] (acute vs chronic)	0.877	.38	0.672	1.382	2.869
Privacy information (PFS[d] vs no PFS)	3.298	<.001	1.888	4.527	11.490
Stigma*privacy information	2.734	.006	1.661	5.952	21.739
TC*privacy information	0.094	.92	0.325	1.058	3.390

[a] R^2=0.107 (Hosmer-Lemeshow), 0.297 (Nagelkerke), 0.284 (Cox-Snell). Model χ^2_5=37.68; P<.001.

[b] OR: odds ratio.

[c] TC: time course.

[d] PFS: privacy fact sheet.

Robustness Check

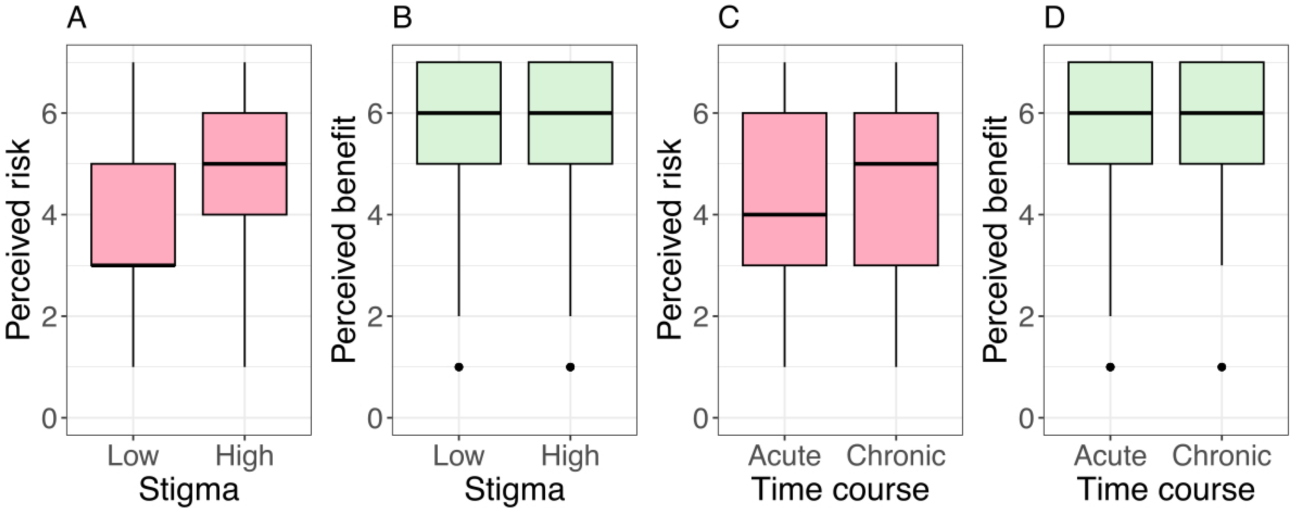
When controlling for interindividual variables (age, gender, education, and mHealth experience), the effects of stigma (z=4.820; P<.001) and information transparency (z=3.548; P<.001) and their interaction (z=3.086; P=.002) remained robust. Age had a negative effect on the upload behavior (z=2.531; OR 0.965, 95% CI 0.939 - 0.992; P=.01). With an increase in age, users were less likely to upload medical findings into their EHR. The other control variables did not influence the upload behavior.

Discussion

Principal Findings

The results of our study show that the decision to upload an individual medical report to an EHR click dummy is influenced by disease-related stigma as well as by privacy notices, that is, concise but comprehensive information about data privacy choices and security measures in the form of PFS. As in our preliminary studies [18,19], uploading diseases with high stigma was associated with increased privacy risk perceptions compared to diseases with low stigma (see Figure 6A), which increased the likelihood of rejecting uploads for stigmatized diseases 6 fold compared to nonstigmatized diseases (see Figure 4A). This finding is surprising given a generally high rating of potential benefits of uploading reports to the EHR (see Figure 6B D) and an overall high willingness to upload medical findings to the EHR. In addition to the nonsignificant manipulation check regarding perceived benefits, uploads did not vary with the TC of the disease (see Figure 4).

Figure 6. Perceived risk as a function of disease-related (A) stigma and (C) time course, and perceived benefit as a function of (B) stigma and (D) time course. The horizontal line in the box represents the median.



Furthermore, a PFS positively influenced the decision to upload. When a PFS was displayed, the likelihood of uploading medical findings to the EHR was more than 3 times higher than when it was not given (see Figure 4C). This is in line with the findings of various studies showing that effective communication of data privacy choices and security information enables people to make informed decisions, thereby reducing general privacy concerns and increasing the use of EHRs [7,24,33,35,61,62]. A plausible explanation for this effect is that a PFS increases users' perceived control over personal data—allowing them to understand and oversee how their health data are handled—which in turn reduces perceived risks and facilitates the decision to upload their data [7,38,40].

Displaying a PFS did not influence the decision to upload medical findings for nonstigmatized diseases, as nearly all nonstigmatized medical reports were uploaded regardless of whether a PFS was given (see Figure 5A). In contrast, for stigmatized diseases, the PFS significantly increased the likelihood of uploads (see Figure 5B). This suggests that showing a PFS shortly before a decision to upload medical findings to the EHR must be made not only is effective in mitigating general privacy concerns but also helps to reduce specific fears related to stigmatized diseases and increase upload decisions [36,37,40].

More generally, studies in nonmedical domains, involving low-risk scenarios such as a shopping assistant [43] and an event finder [36], showed that transparency features or the transparency of privacy policies had no effect on behavior, for instance, on the decision to access the location [43] or the intention to disclose personal data to the event finder [36]. Our findings help to explain these differing findings by showing that the relevance of transparent privacy notes is mainly contingent upon the level of perceived risk associated with the data. In low-risk scenarios, such as nonstigmatized diseases, privacy concerns are typically low, which means that transparency features cannot meaningfully change the decision because there are no substantial concerns to alleviate. However, in high-risk scenarios, such as those involving stigmatized health conditions, privacy concerns are more likely to arise. In these situations, transparency features can help strengthen users' perceived control, reducing perceived risks and thereby increasing acceptance. This highlights the importance of situational context for transparency measures to matter. Transparent information about data privacy and security is not necessary in low-risk situations (although it does not hurt behavioral outcomes), but it becomes crucial for decision making in high-risk contexts, such as the handling of sensitive health data in EHRs.

Implications

The opportunities offered by implementing transparency features in the EHR should be considered by health care stakeholders. Transparency features can not only reduce general privacy concerns but can also address situational concerns triggered by disease-related stigma [18]. Thus, transparency features can ultimately help to ensure equal access to EHRs, even for users who suffer from stigmatized diseases, thereby promoting health equity [49,50,63]. This way, more patients get a chance to benefit from the EHR and, as their illnesses, allergies, and medications can be considered for future diagnostics and therapies, receive better and more targeted treatment.

Limitations and Future Directions

There are several limitations in our study, which need to be considered in subsequent studies. While our manipulation checks for perceived risk (related to stigma) were successful, the manipulation of perceived benefit (related to TC) was not. This may be due to the between-subjects design of our study. In a previous within-subjects design, where participants evaluated both acute and chronic reports, the TC significantly impacted upload behavior [19]. It seems that participants, when comparing multiple conditions, can better discern when uploading is more or less beneficial. In our study, however, participants may have perceived the benefits of uploading as uniformly high, regardless of TC, leading to a diminished ability to detect differences.

It is clear that the adoption and approval of data-gathering technologies are strongly influenced by cultural differences [64]. In comparison to other European nations, the German population exhibits a heightened level of caution regarding the use of personal information online [65]. Given that in this study, data collection was conducted solely with residents of Germany, future studies should validate the applicability of these findings in other countries.

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 disease is not immediately apparent [45,66]. The question arises to what extent the behavior of stigmatized individuals can be simulated under experimental conditions. To further strengthen the validity and generalizability of our results, a follow-up study should examine the perspective of already affected individuals.

Although we captured actual click behavior, the upload decision occurred in a simulated EHR environment using fictitious diagnoses and did not involve participants' actual health data. As such, the behavior measured in this study reflects a behavioral proxy rather than real-world EHR upload behavior. Future work should examine the robustness of these findings and investigate how patients behave when real data, real accounts, and real consequences are involved.

Moreover, this was a survey study with limited immersion despite the use of an interactive click dummy. In a follow-up study, researchers could collaborate with health insurers to gather real-world data on upload behavior with a real EHR and an integrated transparency feature as used for this study. Conversely, our study faced limitations due to uncontrolled conditions like participants' locations and potential distractions, as participants completed the questionnaire online. Future research could validate our findings through a laboratory study, ensuring a more controlled environment.

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 [67,68]. In particular, the level of education of our sample was above average. Although we were unable to detect any effects of the control variables, 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

Our results show that although general upload rates to the EHR are high, stigmatized diseases—even if simulated—negatively affect simulated upload behavior. However, displaying a transparency feature in the form of a PFS increases the likelihood that people upload stigmatized health data when interacting with an EHR click dummy by mitigating privacy concerns. Our findings indicate that the role of transparency features is contingent upon the level of perceived risk associated with the data to be uploaded. When the perceived risk is low, users do not need detailed privacy information to trust the technology and upload their data. However, when uploads involve sensitive data and are seen as risky, users consider privacy information and modify their upload behavior based on the information provided, potentially because transparency features help strengthen users' perceived control in such high-risk situations. Implementing transparency features in EHRs may thus help to ensure that users who perceive high privacy risks when confronted with sensitive health information are not excluded from the benefits of these systems due to privacy concerns, thereby promoting digital health equity.

Acknowledgments

The authors acknowledge support from the German Research Foundation and the Open Access Publication Fund of TU Berlin. They also thank the Evangelisches Studienwerk Villigst and the German Federal Ministry of Education and Research, which provided the doctoral scholarship (NVK) without which this research would not have been possible. They extend their gratitude to all participants of the study for their invaluable contributions.

Funding

This research was supported by the German Research Foundation and the Open Access Publication Fund of Technische Universität Berlin, which covered the publication fees. The doctoral research of NVK was supported by the Evangelisches Studienwerk Villigst and the German Federal Ministry of Education and Research.

Data Availability

The dataset analyzed in this study is available upon reasonable request. Requests to access these datasets should be directed to niklas.vkalckreuth@tu-berlin.de.

Authors' Contributions

NVK researched the literature and conceived the study under the supervision of MAF. NVK was involved in protocol development, gaining ethical approval, participant recruitment, and data analysis. NVK wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript. No generative artificial intelligence tools were used in the conceptualization, data analysis, interpretation, or drafting of the scientific content of this manuscript. Only standard language-editing tools were applied.

Conflicts of Interest

NvK is a member of the JMIR Human Factors editorial board. He was not involved in the editorial or peer-review process of this manuscript. No other conflicts of interest are declared.

Multimedia Appendix 1
Case vignettes.
[PDF File, 96 KB - humanfactors_v13i1e71124_app1.pdf]

Multimedia Appendix 2
Disease descriptions.
[PDF File, 92 KB - humanfactors_v13i1e71124_app2.pdf]

Multimedia Appendix 3
Privacy fact sheets.
[PDF File, 88 KB - humanfactors_v13i1e71124_app3.pdf]

Multimedia Appendix 4
Questionnaire.
[PDF File, 77 KB - humanfactors_v13i1e71124_app4.pdf]

Multimedia Appendix 5
Demographic data of all groups.
[PDF File, 79 KB - humanfactors_v13i1e71124_app5.pdf]

Checklist 1
CONSORT-EHEALTH (v 1.6.1) checklist.
[PDF File, 1352 KB - humanfactors_v13i1e71124_app6.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 and Cyber-Physical Systems 2022;2:12-30. [doi: 10.1016/j.iotcps.2022.04.001]
2. Hubmann M, Pätzmann-Sietas B, Morbach H. Telemedizin und digitale akte – wo stehen wir?: chancen und herausforderungen bei der umsetzung in klinik- und praxisalltag [Article in German]. Monatsschr Kinderheilkd 2021 Aug;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 ASO, Binder S, Amelung VE. Einführung einer elektronischen patientenakte in deutschland vor dem hintergrund der internationalen erfahrungen. In: Klauber J, Geraedts M, Friedrich J, Wasem J, editors. Krankenhaus-Report 2019: Das Digitale Krankenhaus [Book in German]: Springer Berlin Heidelberg; 2019:3-16. [doi: 10.1007/978-3-662-58225-1_1]
5. Baron von Osthoff M, Watzlaw-Schmidt U, Lehmann T, Hübner J. Patientengruppenspezifische datenhoheitsbedürfnisse und akzeptanz der elektronischen patientenakte [Article in German]. Bundesgesundheitsbl 2022 Nov;65(11):1197-1203. [doi: 10.1007/s00103-022-03589-w]
6. Die elektronische patientenakte (epa) für alle [Article in German]. Bundesministerium für Gesundheit. 2021. URL: https://www.bundesgesundheitsministerium.de/elektronische-patientenakte.html [accessed 2026-01-03]
7. 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: Springer International Publishing; 2016:19-50. [doi: 10.1007/978-3-319-23294-2_2]
8. Kus K, Kajüter P, Arlinghaus T, Teuteberg F. Die elektronische Patientenakte als zentraler bestandteil der digitalen transformation im Deutschen gesundheitswesen – eine analyse von akzeptanzfaktoren aus patientensicht [Article in German]. HMD 2022 Dec;59(6):1577-1593. [doi: 10.1365/s40702-022-00921-5]
9. Verband der Ersatzkassen. Daten zum gesundheitswesen: versicherte [Article in German]. vdek. 2024. URL: https://www.vdek.com/presse/daten/b_versicherte.html [accessed 2026-01-03]

10. Lauterbach: elektronische patientenakte ab ende 2024 für alle verbindlich [Article in German]. Bundesministerium für Gesundheit. 2023. URL: https://www.bundesgesundheitsministerium.de/presse/interviews/interview/fas-030324-elektronische-patientenakte.html [accessed 2026-01-03]
11. Die elektronische patientenakte (epa) erleichtert den alltag [Article in German]. Bundesministerium für Gesundheit. 2025. URL: https://www.bundesgesundheitsministerium.de/epa-na-sicher/ [accessed 2026-01-03]
12. Hoerbst A, Kohl CD, Knaup P, Ammenwerth E. Attitudes and behaviors related to the introduction of electronic health records among Austrian and German citizens. Int J Med Inform 2010 Feb;79(2):81-89. [doi: 10.1016/j.ijmedinf.2009.11.002] [Medline: 20031482]
13. Kröner S, Schreiweis B, Strotbaum V, Brandl LC, Pobiruchin M, Wiesner M. Consumer perspectives on the national electronic health record and barriers to its adoption in Germany: does health policy require a change in communication? BMC Health Serv Res 2025 Jan 7;25(1):33. [doi: 10.1186/s12913-024-12175-6] [Medline: 39773635]
14. Panskus R, Bochum RU, Ninow M, Fahl S, Marky K, Bochum RU. Privacy mental models of electronic health records: a German case study. Presented at: Proceedings of the Nineteenth Symposium on Usable Privacy and Security; Aug 7-8, 2023 URL: https://www.usenix.org/system/files/soups2023-panskus.pdf [accessed 2026-01-03]
15. Papoutsi C, Reed JE, Marston C, Lewis R, Majeed A, Bell D. Patient and public views about the security and privacy of Electronic Health Records (EHRs) in the UK: results from a mixed methods study. BMC Med Inform Decis Mak 2015 Oct 14;15(1):86. [doi: 10.1186/s12911-015-0202-2] [Medline: 26466787]
16. “Konnte bisher noch nie gehackt werden”: die elektronische patientenakte kommt - jetzt für alle! 2024 [Article in German]. Media.ccc.de. URL: https://media.ccc.de/v/38c3-konnte-bisher-noch-nie-gehackt-werden-die-elektronische-patientenakte-kommt-jetzt-fr-alle [accessed 2026-01-03]
17. Stellungnahme zum CCC-vortrag zur epa für alle [Article in German]. Gematik. 2024 Dec 27. URL: https://www.gematik.de/newsroom/news-detail/aktuelles-stellungnahme-zum-ccc-vortrag-zur-epa-fuer-alle [accessed 2026-01-03]
18. 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 Apr 10;11:e52625. [doi: 10.2196/52625] [Medline: 38598271]
19. 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. Presented at: Americas Conference on Information Systems (AMCIS 2023) Proceedings: Diving into Uncharted Waters; Aug 10-12, 2023 URL: https://aisel.aisnet.org/amcis2023/sig_health/sig_health/2/ [accessed 2026-01-03]
20. von Kalckreuth N, Feufel MA. Disease characteristics influence the privacy calculus to adopt electronic health records: a survey study in Germany. Digit Health 2024;10:20552076241274245. [doi: 10.1177/20552076241274245] [Medline: 39247096]
21. Rau E, Tischendorf T, Mitzscherlich B. Implementation of the electronic health record in the German healthcare system: an assessment of the current status and future development perspectives considering the potentials of health data utilisation by representatives of different stakeholder groups. Front Health Serv 2024;4:1370759. [doi: 10.3389/frhs.2024.1370759] [Medline: 38800500]
22. Feng Y, Yao Y, Sadeh N. A design space for privacy choices: towards meaningful privacy control in the internet of things. Presented at: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan p. 1-16. [doi: 10.1145/3411764.3445148]
23. Sloan RH, Warner R. Beyond notice and choice: privacy, norms, and consent. SSRN Journal 2013;14(2). [doi: 10.2139/ssrn.2239099]
24. Kelley PG, Cranor LF, Sadeh N. Privacy as part of the app decision-making process. : Association for Computing Machinery; 2013 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 27 to May 2, 2013; Paris France p. 3393-3402. [doi: 10.1145/2470654.2466466]
25. Schaub F, Könings B, Weber M. Context-adaptive privacy: leveraging context awareness to support privacy decision making. IEEE Pervasive Comput 2015 Jan;14(1):34-43. [doi: 10.1109/MPRV.2015.5]
26. Meier Y, Schäwel J, Krämer NC. The shorter the better? Effects of privacy policy length on online privacy decision-making. MaC 2020 Jun 23;8(2):291-301. [doi: 10.17645/mac.v8i2.2846]
27. Fabian B, Ermakova T, Lentz T. Large-scale readability analysis of privacy policies. Presented at: Proceedings of the International Conference on Web Intelligence; Aug 23-26, 2017; Leipzig, Germany p. 18-25. [doi: 10.1145/3106426.3106427]
28. McDonald AM, Cranor LF. The cost of reading privacy policies. I/S: J Law Policy Inf Soc 2008;4(3):543-568 [FREE Full text]
29. Reidenberg JR, Breaux T, Cranor LF, et al. Disagreeable privacy policies: mismatches between meaning and users understanding. SSRN J 2015;30. [doi: 10.2139/ssrn.2418297]
30. Obar JA, Oeldorf-Hirsch A. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. Inf Commun Soc 2020 Jan 2;23(1):128-147. [doi: 10.1080/1369118X.2018.1486870]
31. Braghin C, Cimato S, Della Libera A. Are mHealth apps secure: case study. Presented at: 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC); Jul 23-27, 2018; Tokyo, Japan p. 335-340. [doi: 10.1109/COMPSAC.2018.10253]

32. van Haasteren A, Gille F, Fadda M, Vayena E. Development of the mHealth App Trustworthiness checklist. Digit Health 2019;5:2055207619886463. [doi: 10.1177/2055207619886463] [Medline: 31803490]
33. Kulyk O, Milanovic K, Pitt J. Does my smart device provider care about my privacy? Investigating trust factors and user attitudes in IoT systems. Presented at: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society; Oct 25-29, 2020; Tallinn, Estonia p. 1-12. [doi: 10.1145/3419249.3420108]
34. Krasnova H, Spiekermann S, Koroleva K, Hildebrand T. Online social networks: why we disclose. J Inf Technol 2010 Jun;25(2):109-125. [doi: 10.1057/jit.2010.6]
35. 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, 1st edition: Cambridge University Press; 2018:365-393. [doi: 10.1017/9781316831960.021]
36. Karwatzki S, Dytynko O, Trenz M, Veit D. Beyond the personalization–privacy paradox: privacy valuation, transparency features, and service personalization. J Manage Inf Syst 2017 Apr 3;34(2):369-400. [doi: 10.1080/07421222.2017.1334467]
37. Tsai JY, Egelman S, Cranor L, Acquisti A. The effect of online privacy information on purchasing behavior: an experimental study. Inf Syst Res 2011 Jun;22(2):254-268. [doi: 10.1287/isre.1090.0260]
38. 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. [doi: 10.2196/45503] [Medline: 37585259]
39. 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]
40. von Kalckreuth N, Kopka M, Appel J, Feufel MA. Unlocking the potential of the electronic health record - the influence of transparency features. Presented at: Proceedings of the 32nd European Conference on Information Systems (ECIS 2024); Jun 13-19, 2024; Paphos, Cyprus p. 1-15 URL: https://aisel.aisnet.org/ecis2024/track18_healthit/track18_healthit/5/ [accessed 2026-01-03]
41. von Kalckreuth N, Feufel MA. Disclosure of health data – conceptualizing the intention to use wearables as an extended privacy calculus. Presented at: Proceedings of the Americas Conference on Information Systems (AMCIS 2021); Aug 9-13, 2021; Montreal, Canada p. 1-6 URL: https://aisel.aisnet.org/amcis2021/sig_hci/sig_hci/6/ [accessed 2026-01-03]
42. Adjerid I, Acquisti A, Brandimarte L, Loewenstein G. Sleights of privacy: framing, disclosures, and the limits of transparency. Presented at: Proceedings of the Ninth Symposium on Usable Privacy and Security; Jul 24-26, 2013; Newcastle, United Kingdom p. 1-11. [doi: 10.1145/2501604.2501613]
43. Betzing JH, Tietz M, vom Brocke J, Becker J. The impact of transparency on mobile privacy decision making. Electron Markets 2020 Sep;30(3):607-625. [doi: 10.1007/s12525-019-00332-3]
44. von Kalckreuth N, Feufel MA. Enhancing uploads of health data in the electronic health record-the role of framing and length of privacy information: a survey study in Germany. Digit Health 2025;11:20552076251351197. [doi: 10.1177/20552076251351197] [Medline: 40755954]
45. Goffman E. Stigma: Notes on the Management of Spoiled Identity: Prentice-Hall, Inc; 1963. URL: https://cdn.penguin.co.uk/dam-assets/books/9780241548011/9780241548011-sample.pdf [accessed 2026-01-03]
46. 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]
47. Keith MJ, Thompson SC, Hale J, Lowry PB, Greer C. Information disclosure on mobile devices: re-examining privacy calculus with actual user behavior. Int J Hum Comput Stud 2013 Dec;71(12):1163-1173. [doi: 10.1016/j.ijhcs.2013.08.016]
48. Keith MJ, Clark A, Masters T, Wigington C. What makes health data privacy calculus unique? Separating probability from impact. Presented at: Proceedings of the 55th Hawaii International Conference on System Sciences; Jan 4-7, 2022; Maui, HI p. 4836-4845. [doi: 10.24251/HICSS.2022.589]
49. Stablein T, Hall JL, Pervis C, Anthony DL. Negotiating stigma in health care: disclosure and the role of electronic health records. Health Sociol Rev 2015 Sep 2;24(3):227-241. [doi: 10.1080/14461242.2015.1078218]
50. Link BG, Phelan JC. Conceptualizing stigma. Annu Rev Sociol 2001 Aug;27(1):363-385 [FREE Full text] [doi: 10.1146/annurev.soc.27.1.363]
51. Cheng YH, Yeh YJ. Exploring radio frequency identification technology's application in international distribution centers and adoption rate forecasting. Technol Forecast Soc Change 2011 May;78(4):661-673. [doi: 10.1016/j.techfore.2010.10.003]
52. Lancelot Miltgen C, Popovič A, Oliveira T. Determinants of end-user acceptance of biometrics: integrating the "Big 3" of technology acceptance with privacy context. Decis Support Syst 2013 Dec;56:103-114. [doi: 10.1016/j.dss.2013.05.010]
53. Bickford J, Barton SE, Mandalia S. Chronic genital herpes and disclosure.... the influence of stigma. Int J STD AIDS 2007 Sep;18(9):589-592. [doi: 10.1258/095646207781568484] [Medline: 17784999]
54. Zacks S, Beavers K, Theodore D, 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]
55. Berger BE, Ferrans CE, Lashley FR. Measuring stigma in people with HIV: psychometric assessment of the HIV stigma scale. Res Nurs Health 2001 Dec;24(6):518-529. [doi: 10.1002/nur.10011] [Medline: 11746080]
56. Zarei N, Joulaei H, Darabi E, Fararouei M. Stigmatized attitude of healthcare providers: a barrier for delivering health services to HIV positive patients. Int J Community Based Nurs Midwifery 2015 Oct;3(4):292-300. [Medline: 26448956]

57. Schomakers EM, 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. [doi: 10.2196/27095] [Medline: 35040801]
58. 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. [doi: 10.1007/s10916-023-01910-x] [Medline: 36705853]
59. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
60. Cuschieri S. The CONSORT statement. Saudi J Anaesth 2019 Apr;13(Suppl 1):S27-S30. [doi: 10.4103/sja.SJA_559_18] [Medline: 30930716]
61. 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]
62. Kroll T, Stieglitz S. Digital nudging and privacy: improving decisions about self-disclosure in social networks. Behav Inf Technol 2021 Jan 2;40(1):1-19. [doi: 10.1080/0144929X.2019.1584644]
63. Hatzenbuehler ML, Phelan JC, Link BG. Stigma as a fundamental cause of population health inequalities. Am J Public Health 2013 May;103(5):813-821. [doi: 10.2105/AJPH.2012.301069] [Medline: 23488505]
64. Krasnova H, Veltri NF, Günther O. Self-disclosure and privacy calculus on social networking sites: the role of culture: intercultural dynamics of privacy calculus. Bus Inf Syst Eng 2012 Jun;4(3):127-135. [doi: 10.1007/s12599-012-0216-6]
65. Custers B, Dechesne F, Sears AM, Tani T, van der Hof S. A comparison of data protection legislation and policies across the EU. Comput Law Secur Rev 2018 Apr;34(2):234-243. [doi: 10.1016/j.clsr.2017.09.001]
66. 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]
67. Bildungsstand [Article in German]. Statistisches Bundesamt. 2020. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Bildungsstand/_inhalt.html [accessed 2026-01-03]
68. Bevölkerungsstand [Article in German]. Statistisches Bundesamt. 2022. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/_inhalt.html [accessed 2026-01-03]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
EHR: electronic health record
IV: independent variable
mHealth: mobile health
OR : odds ratio
PFS: privacy fact sheet
SP: stigmatization potential
TC: time course

Edited by J Sarvestan; submitted 10.Jan.2025; peer-reviewed by M Greve, Y Senathirajah; revised version received 17.Nov.2025; accepted 18.Nov.2025; published 15.Jan.2026.

Please cite as:
von Kalckreuth N, Feufel MA
Privacy Fact Sheets for Mitigating Disease-Related Privacy Concerns and Facilitating Equal Access to the Electronic Health Record: Randomized Controlled Trial
JMIR Hum Factors 2026;13:e71124
URL: https://humanfactors.jmir.org/2026/1/e71124
doi:10.2196/71124

Exploring Medical Information Needs and Accessibility in Swedish Dental Care by Analysis of Documentation Workflows and Electronic Dental Records in Dalarna: Sociotechnical Qualitative Study

Sahid Hasan Rahim, MSc, DDS; Nadia Davoody, MSc, PhD; Stefano Bonacina, MSc, PhD

Health Informatics Centre, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Tomtebodavägen 18 A, Stockholm, Sweden

Corresponding Author:
Stefano Bonacina, MSc, PhD
Health Informatics Centre, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Tomtebodavägen 18 A, Stockholm, Sweden

Abstract

Background: Despite growing evidence demonstrating the connection between oral and systemic health, medical and dental care remain institutionally divided. A significant consequence of this division is the lack of information sharing, which is particularly problematic in dental care, where knowing patients' medical information is crucial for providing safe and effective treatments. This separation poses additional challenges in Swedish regions with limited resources, such as Dalarna, where dental care practices would benefit from improved access to relevant medical information in their electronic dental record (EDR) systems.

Objective: This study aimed to explore how current documentation workflows and EDR systems support the medical information needs within dental care practices in Dalarna and consider what influence direct access to medical information could have.

Methods: The study adopted an exploratory-descriptive qualitative approach. Semi-structured interviews were conducted with dental practitioners working in general dental practices. Data collection followed a sociotechnical framework, and thematic analysis was performed to identify key medical information needs, as well as current workflow and system limitations. Conceptual models were developed to reflect these findings.

Results: Eighteen dental practitioners were interviewed. The identified medical information needs included specific types of medical conditions, pharmacological information, treatment history, and laboratory values. Furthermore, dental practitioners highlighted substantial challenges in existing documentation workflows and the EDR system. Proposed conceptual models demonstrated how integrating EDR systems with the Swedish National Patient Overview ("Nationell Patientöversikt") via National Service Platform ("Nationell Tjänsteplattform") could streamline workflows and enhance information accessibility.

Conclusions: The findings show a clear need to improve medical information accessibility in dental care. A solution is to facilitate interoperability and align digital infrastructure with the identified needs. The proposed recommendations offer a feasible starting point for improving medical information access in Swedish dental care, particularly in resource-constrained regions like Dalarna.

(JMIR Hum Factors 2026;13:e82691) doi:10.2196/82691

KEYWORDS

dental informatics; dental care; health information exchange; health information interoperability; health information systems; medical informatics

Introduction

Medical Information in Dental Care

Comprehensive medical information is essential in dental care, as inaccurate or incomplete patient medical history assessments can lead to adverse treatment outcomes [1]. In addition to patient safety, the absence of accurate medical information exposes dental professionals to potential harm, as many procedures are invasive, carrying risks of occupational hazards such as cross-contamination and disease transmission [2]. The literature highlights several types of medical information as particularly relevant to dental care. Medical history, especially patients' diagnoses and current medications, is consistently prioritized [3-5]. Information about allergies and adverse drug reactions is also essential, as overlooking these may lead to severe or even life-threatening scenarios [3,4,6]. Additionally, details such as medical treatment plans, hospitalization history, and surgical history are considered highly valuable, as they support improved care coordination and more effective treatment planning [3-5]. Together, these categories form the foundation of the medical information that dental professionals rely on to deliver care. Unfortunately, challenges arise due to inconsistencies in the documentation and communication of such information, particularly as it is often self-reported by patients [7]. This raises the question of whether health information technology can possibly address this gap and support more reliable access to critical medical information in dental care.

Information Technology for Record-Keeping

In dental care, patient information is primarily managed through electronic dental records (EDRs), whereas medical care relies on electronic health records (EHRs) [8-10]. The advent of such information systems has facilitated the process of recording patient information in data format, thereby giving professionals in both dental and medical care a more structured and streamlined approach to record-keeping [9,10]. This also includes the process of enabling information exchange.

Interoperability

To enable the exchange of patient information between EHR and EDR systems, interoperability is required [11]. Interoperability refers to the ability of different systems to exchange, process, and use patient data effectively, and it is commonly described across four levels. The technical level involves a shared technical infrastructure that allows systems to communicate. The syntactic level requires data to be structured in a standardized format for processing. The semantic level ensures that data are interpretable and meaningful across systems. Finally, the organizational level emphasizes that interoperability at the technical, syntactic, and semantic levels depends on alignment with regulatory, administrative, and policy frameworks, including agreements that enable data exchange while ensuring compliance with laws and regulations.

Current Dental-Medical Interoperability Efforts

The literature identifies two main approaches to achieving interoperability between dental and medical care systems. Health information exchange (HIE) serves as an intermediary to facilitate data exchange between different systems [3,12-14]. Integrated EHR-EDR (IEHR), on the other hand, operates on a single-system architecture [14]. This system has become widely adopted by organizations offering integrated dental-medical care services [15].

While both IEHR and HIE aim to facilitate information exchange between dental and medical care, their feasibility depends on organizational size, infrastructure, and financial resources. IEHR can provide data conformity and reliability, but may be financially and technically restrictive for smaller organizations [4,7,14]. In contrast, HIE can offer a more flexible and cost-effective solution, though usability barriers may impede its adoption [3,12-14].

Understanding how these approaches are implemented within a specific national context is essential for assessing their real-world applicability. Here, Sweden offers a particularly instructive case due to its publicly funded health care system, high digital maturity, and ongoing efforts to advance HIE [16,17]. In fact, Sweden's broader aim is to become a global leader in health care digitalization by 2025, a vision that includes both medical and dental care [18].

The Swedish Interoperability Strategy

Sweden is a country divided into 21 administrative regions, each governing its publicly funded medical and dental care services autonomously [19]. Due to this decentralized administrative structure, challenges have arisen in ensuring the accessibility and exchange of patient data across regions and organizations [17,20]. Hence, the Swedish government enacted the Integrated Health and Social Care Documentation Act (2022:913), establishing a legal basis for the sharing of patient information across health care services [21]. This legal foundation was followed by the implementation of nationwide digital infrastructure, the national patient overview (nationell patientöversikt; NPÖ), and the national service platform (nationell tjänsteplattform; NTjP) [22-24]. NPÖ is a nationwide digital platform primarily provided to public health care providers, allowing authorized health care professionals to retrieve patient data across organizations and regions [23]. NTjP refers to the technical infrastructure that facilitates the exchange of patient data in a standardized and secure manner [22,24]. This concept is similar to HIE, as it acts as an intermediary layer between different systems in organizations and regions.

Currently, the digital infrastructure does not support information exchange between dental and medical care services [25]. It may pose particular challenges in regions where access to both services is already constrained, particularly in underserved or rural regions [26-29]. To illustrate the relevance and urgency of such an approach, the region of Dalarna provides a meaningful context for further discussion.

The Dalarna Region

One of the most pressing issues in Dalarna is the shortage of dental professionals, with the lowest dentist-to-population ratio in Sweden [30]. Due to its vast land area, patients in remote regions may need to travel long distances to reach a general dental practice (GDP). In addition, Dalarna has an aging population, with nearly 25% of residents aged 65 years or older, compared to the national average of 19% [31]. This demographic trend places further strain on an already limited dental workforce, as older adults often require more frequent and complex dental and medical care [32]. Facilitating access to medical information in dental care in Dalarna could therefore support more effective and higher-quality care outcomes for patients.

Problem Description

Sweden's digital infrastructure, the NPÖ and the NTjP, support information exchange between health care information systems across regions and organizations [22-24]. However, these platforms do not currently provide access to medical information in dental care [25]. This lack of integration presents a missed opportunity to enhance both patient safety and efficiency of dental workflows, especially in rural regions where access to medical and dental services is limited [26-29]. The Dalarna region exemplifies this challenge, facing one of the most severe shortages of dental professionals in the country. This shortage forces many residents to travel long distances for dental care [30]. Yet, addressing this gap does not necessarily require extensive structural or financial reforms. Rather, a pragmatic path forward may lie in leveraging existing infrastructure to enable dental professionals to access relevant medical information directly through their EDR systems. However, without a clear understanding of what specific medical information is needed and how it fits into dental workflows, such interoperability efforts risk being ineffective or even disruptive. Although existing literature reports the types of medical information relevant to dental care [3-5], it lacks specificity. Therefore, it is crucial to identify the specific medical information needs of dental professionals and examine how these can be supported within current workflows, systems, and organizational structures. This necessitates a sociotechnical approach that accounts for both technical components and end-users [33,34], primarily dental professionals.

Aim

This study aims to explore how current documentation workflows and EDR systems in dental care practices in the Dalarna region in Sweden support or limit the fulfillment of the medical information needs, and to examine the potential impact of having direct access to accurate and relevant medical information within EDR systems.

Methods

Research Design

This study adopted an exploratory-descriptive qualitative research approach to allow for a comprehensive yet structured and contextualized examination of the experiences, challenges, and needs on the subject matter [35].

Study Setting

The study was decided to be conducted in GDPs operated under the Public Dental Service Dalarna (PDSD), which is administered by the regional authority Region Dalarna and is responsible for providing dental care services within the region [36]. Consequently, private dental clinics or organizations and specialist clinics were excluded.

Participants and Sampling

The primary participants consisted of dental practitioners, as they are directly responsible for patient care [37]. Only dental practitioners having their main professional duties within GDPs operated by PDSD were considered. Purposive sampling was used to identify participants who were able to provide detailed, precise, and reliable information aligned with the study's aim [38]. Conversely, other dental professionals were excluded.

Data Collection

Interviews were deemed the most appropriate data collection method as they offered an opportunity for participants to share their perspectives in-depth related to the subject matter [39]. The interviews followed a semi-structured format to ensure a balance between systematic inquiry and flexibility. The interviews needed to capture the sociotechnical aspects relevant to the subject matter. To achieve this, an interview guide was designed following a sociotechnical framework. The sociotechnical framework defined by Sittig et al [40] provides a comprehensive lens for evaluating health information technology by examining the interaction between information systems and the social context in which they operate. It comprises 8 dimensions, where for this study, 6 dimensions were chosen based on their alignment with the research aim [40]: (1) clinical content to identify the specific medical information considered important and what should be accessible in EDR systems; (2) people to explore experiences and challenges in obtaining medical information from patients, the medical domain, or other sources; (3) human-computer interface to assess the usability of the current EDR system regarding accessibility and documentation of medical information; (4) workflow and communication to understand how medical information is obtained and documented within current workflows, focusing on the patient and EDR system; (5) internal organizational policies, procedures, and culture to examine existing policies on documentation and retrieval of medical information within PDSD; and (6) external rules, regulations, and pressures to explore the awareness of guidelines or requirements for documenting medical information in dental care. These dimensions informed the development of the interview guide in Swedish and English (Multimedia Appendix 1).

Assessment and Pilot Testing

Following development, the interview guide was reviewed by the co-authors, SB, ND, and the Head of Division at PDSD, to confirm its suitability. The guide was subsequently pilot-tested, which confirmed that the questions elicited comprehensive responses. Individuals involved in the review and pilot testing were excluded from the final study to avoid diluting the results. A bilingual consent form in both Swedish and English was developed in parallel (Multimedia Appendix 2).

Recruitment and Interview Sessions

The head of division at PDSD facilitated recruitment by distributing interview invitations to GDPs. Interviews were conducted digitally via Microsoft Teams between March 7 and July 28, 2025. Digital interviews were chosen to accommodate participants from various locations and minimize travel-related costs. All interviews were audio-recorded and transcribed using Microsoft Teams' built-in tool. Supplementary notes were taken during each session to capture contextual and paralinguistic cues. Transcripts were carefully reviewed, compiled, and returned to participants to confirm accuracy. Interviewing concluded when data saturation was reached, meaning no new information emerged, and the data were sufficiently rich to address the study's research questions [41].

Data Analysis

Thematic Analysis

Thematic analysis was conducted to identify and interpret patterns within the interview data, guided by the sociotechnical framework from Sittig et al [40]. The analysis followed Braun and Clarke's 6-phase approach [42]. Transcripts were read repeatedly to gain an in-depth understanding of the data, with audio recordings revisited when meaning was unclear. Excerpts were then collated into potential themes, which were reviewed and refined to ensure they accurately represented the data. Themes were organized under the following columns: sociotechnical dimension, participant (anonymized), excerpt, subtheme, and theme. The entire analysis was iterative, with frequent revisions of subthemes and themes to ensure a valid and coherent representation of the findings [43].

Conceptual Modeling

Following the analysis, conceptual models were developed to illustrate current and proposed workflows for accessing medical information in dental care. These models considered integration with existing Swedish infrastructures. Unified Modeling Language (UML) was used due to its use in health system planning and stakeholder communication [44,45]. Three UML diagrams were created. Activity diagrams were used to show current and improved workflows for obtaining and documenting medical information in dental care. Use case diagrams were developed to present expected system functionalities derived from the interview data. Finally, a sequence diagram was created to illustrate how medical information could be exchanged from regional or national infrastructures to local EDR systems.

Ethical Considerations

The study was carried out in Sweden. According to the Swedish Ethical Review Act (SFS 2003:460) [46] and guidance from the Swedish Ethical Review Authority [47], the type of research presented in this article does not require formal ethical approval, as it does not involve sensitive personal data as defined by the European Union General Data Protection Regulation (EU 2016/679) [48]. Nonetheless, we emphasize that all ethical standards were strictly followed in accordance with relevant legislation, the Declaration of Helsinki [49]. Participants were fully informed about the study's purpose, their rights, and data handling procedures. Informed consent was obtained both in writing and verbally. Participation was voluntary, and withdrawal was permitted at any time. Anonymity was guaranteed, and participants were informed that findings would be publicly disseminated.

Results

Descriptive Overview of the Interview Sample

In total, 18 interviews were conducted, after which no new information emerged beyond the 11th interview. All participants reported using the T4 EDR system, developed by Carestream [50]. Fifteen participants identified as female and three as male. Key characteristics of the participants are presented in Table 1. All participants treated diverse groups of patients and performed a broad range of general dental treatments, where 8 participants identified oral surgery as their main focus. All participants primarily used free text.

Table . Key characteristics of the participants.

Participants	Age (year range)	Work experience overall (year range)	Work experience Dalarna (year range)	Scope of practice
P1	40 - 45	10 - 15	10 - 15	General Dentistry and Oral Surgery
P2	25 - 30	0 - 5	0 - 5	General Dentistry
P3	25 - 30	0 - 5	0 - 5	General Dentistry
P4	25 - 30	0 - 5	0 - 5	General Dentistry, Pediatric Dentistry, and Orthodontics
P5	25 - 30	0 - 5	0 - 5	General Dentistry and Oral Surgery
P6	25 - 30	0 - 5	0 - 5	General Dentistry
P7	45 - 50	20 - 25	10 - 15	General Dentistry, Oral Surgery, Pediatric Dentistry, and Orthodontics
P8	25 - 30	0 - 5	0 - 5	General Dentistry, Oral Surgery, and Prosthodontics
P9	45 - 50	25 - 30	5 - 10	General Dentistry and Oral Surgery
P10	25 - 30	0 - 5	0 - 5	General Dentistry and Oral Surgery
P11	25 - 30	0 - 5	0 - 5	General Dentistry, Oral Surgery, and Pediatric Dentistry
P12	20 - 25	0 - 5	0 - 5	General Dentistry
P13	25 - 30	0 - 5	0 - 5	General Dentistry
P14	35 - 40	5 - 10	0 - 5	General Dentistry, Oral Surgery, and Prosthodontics
P15	35 - 40	0 - 5	0 - 5	General Dentistry
P16	25 - 30	0 - 5	0 - 5	General Dentistry
P17	35 - 40	5 - 10	5 - 10	General Dentistry and Community Dentistry
P18	25 - 30	0 - 5	0 - 5	General Dentistry

The Specific Medical Information Needs – Clinical Content

Under the sociotechnical dimension of “clinical content,” the interview data revealed three overarching categories of medical information needs: “conditions, pharmacological information, treatment history and laboratory values.” Each category contains specific types of information that participants deemed essential for providing dental care (Table 2).

Table . Specific medical information identified within the sociotechnical dimension “clinical content.”

Themes	Participants
Conditions	
Oncological diagnosis	P3, P5, P7, P8, P9, P11, P12, P13, P15, P16, P17, P18
Psychiatric and neurological disorders	P2, P4, P7, P11, P12, P14, P18
Cardiovascular diseases	P2, P3, P5, P8, P13, P14, P15, P16, P17, P18
Hematologic disorder	P1, P15, P17
Hepatic diseases	P3
Pulmonary diseases	P3, P17
Renal diseases	P3, P17, P18
Rheumatism	P3
Diabetes	P7, P14
Sleep apnea	P9
Infectious disease	P11, P17
Allergies in general	P2, P12, P14, P15, P18
Pharmacological information	
Penicillin allergy	P11, P14, P16
Substance abuse	P11
Comprehensive medication list	P1, P3, P4, P6, P10, P12, P15
Anticoagulants	P1, P2, P5, P6, P7, P8, P13, P14, P16, P17, P18
Antiresorptive agents	P1, P2, P3, P6, P7, P8, P9, P10, P11, P12, P13, P14, P16, P17, P18
Psychiatric medications	P2, P4, P11, P12, P16, P17, P18
Treatment history and laboratory values	
Ongoing oncological treatment	P5, P7, P8, P9, P11, P12, P16, P17
Surgical procedures	P1, P5, P10, P13, P18
Blood coagulation marker	P8, P11, P14

Conditions

Among conditions, oncological diagnoses and cardiovascular diseases were the most frequently emphasized because of the associated risks when providing invasive treatments. For patients with oncological diagnoses, participants highlighted that treatment planning and risk management are directly influenced by the potential effects of cancer on oral healing and immune response. Similarly, participants reported that patients with cardiovascular diseases require special attention due to systemic risks associated with invasive procedures:

> *Knowing if a patient has cancer or has had it in the past can influence the treatment we provide in terms of healing outcomes […] healing can be delayed, and there is a greater risk for them to get infected.* [P9]
>
> *To be aware of patients with cardiovascular diseases, where the choice of anesthesia matters. There is also a bleeding risk, and some may need antibiotic prophylaxis.* [P14]

Psychiatric and neurological disorders were also commonly mentioned due to their impact on patient cooperation and communication:

> *Neurodivergent conditions matter, particularly when working with children. It influences how you communicate and treat them.* [P4]

Allergies in general were also deemed important, as considerations need to be taken in relation to the materials used in dental care:

> *Some patients have sensitivities to certain materials, so we need to be careful. This is particularly important to know in [dental] prosthetic treatments, where you are dealing with crowns or removable dentures.* [P14]

Additionally, infectious diseases were highlighted due to their relevance in maintaining a safe clinical environment for both patients and dental professionals:

> *Being aware of any infectious diseases is important, not only for our own safety but also to prevent cross-contamination between patients.* [P11]

Other conditions were brought up by individual participants. Though each was mentioned less frequently, they reflect the complexity of the oral-systemic health correlation in the context of coordinating dental care:

> *Patients often have several different diagnoses, such as heart problems, kidney issues, liver disease, or lung conditions, and each of those can have a real impact on oral health. Sometimes it is not clear how it all is related [for the patient], but it definitely affects how we manage them chairside.* [P3]
>
> *For patients with diabetes, wound healing is affected, and there is a higher risk of infection and periodontal disease.* [P14]

Pharmacological Information

The pharmacological information theme comprises medications that participants regarded as critical, primarily due to the procedural risks associated with specific drugs. The most frequently mentioned medication category was antiresorptive agents and anticoagulants due to complications associated with dental treatments:

> *If we miss that someone has been on multiple blood thinners or bisphosphonates, and they do not mention it because they're no longer taking them or do not think it is important [to mention], it can lead to problems like excessive bleeding or delayed wound healing.* [P13]

Additionally, several participants expressed the importance of knowing whether patients are taking psychiatric medications, since that would increase their understanding of patient behavior and anxiety-induced responses during care:

> *I would actually like to see more information about psychiatric medications and prescriptions. I feel like we often miss that type of medication. We could better tailor our care for these patients, particularly in cases where these medications, combined with anesthetics or sedatives, affect how a patient responds to treatment, like increased anxiety.* [P2]

Penicillin allergies and substance abuse were also mentioned, as there were concerns about prescribing antibiotics or analgesics, particularly due to risks of adverse reactions or potential misuse by patients:

> *Someone who lived in another region started coming to clinics here in Dalarna asking for morphine. We did not know about the substance abuse at first, but then we received a warning from the local pharmacy saying we should not prescribe anything to her. […] If the patient is allergic to penicillin, we need to know so we do not prescribe it.* [P11]

Treatment History and Laboratory Values

In treatment history and laboratory values, participants frequently considered ongoing oncological treatment as essential information for treatment planning and to prevent complications:

> *If they [patients] undergo radiation therapy [for cancer], it is really important that we know about it. Not just that they had treatment, but also which region was irradiated. That kind of detail affects how we approach even routine dental procedures.* [P9]

Surgical procedures were also referenced since such information is used to assess potential risks and determine whether treatment should be delayed:

> *If someone has had an operation, we need to know, because they may have to receive antibiotic prophylaxis. I may also decide to postpone treatments.* [P18]

Likewise, information on blood coagulation markers was commonly sought to support decisions around the timing of treatments for patients and to evaluate their bleeding risks:

> *For patients on warfarin, having access to their International Normalized Ratio value is essential. It helps me determine whether I can proceed with treatment immediately or if I need to postpone or take extra precautions.* [P8]

Current Workflow: People

The sociotechnical dimension of people focused on identifying who the participants interact with in relation to obtaining medical information. As summarized in Table 3, it revealed that medical information is predominantly derived from patient reporting, proxy reporting, or through medical consultation. In general, the findings demonstrated how human factors influence the workflow, suggesting the need for medical information to be made more accessible.

Table . Sources for obtaining medical information, categorized under the sociotechnical dimension of "People".

Themes	Participants
Patient reporting	
Verbal recall	P1, P2, P3, P4, P6, P8, P9, P12, P14, P15, P16, P17, P18
Medication list	P5, P7, P8, P10, P11, P13, P15, P16, P17, P18
Online patient portal	P2, P8, P10, P11, P13, P18
Incomplete or inaccurate reporting	P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P13, P14, P15, P16, P17, P18
Proxy reporting	
Companions	P1, P2, P17
Care home	P1, P2, P5, P12
Direct relatives	P5, P11, P17
Challenges in proxy reporting	P1, P2, P11, P12
Medical consultations	
Primary care	P1, P2, P4, P5, P6, P8, P10, P11, P13, P15, P17
Hematology	P5, P7, P9, P12
Cardiology	P9, P11, P18
Oncology	P1, P9
Oral and Maxillofacial Surgery	P1, P10
Nephrology	P9
Coordination limitations	P1, P2, P4, P5, P6, P7, P8, P10, P13, P14, P15

Patient Reporting

Participants reported that medical information was often obtained through verbal recall, medication lists, or by asking patients to access their online patient portal. Verbal recall relied on patients remembering their medical history; medication lists were paper documents brought in by patients; and the online patient portal was used to verify details when verbal reports were insufficient, or no list was available:

> *Medical history is mostly obtained from the patients themselves verbally.* [P4]

> *Sometimes I ask the patient to log in to their online patient portal. The elderly do not know how, so I usually rely on what they say, or ask them to bring a medication list later.* [P8]

A recurring issue was incomplete or inaccurate reporting, mainly due to patients forgetting diagnoses or medications, particularly amongst the elderly or those without a caregiver:

> *Older patients may not know their diagnoses or medications. Also, some patients deliberately leave things out. Just asking is not always enough.* [P4]

Participants also noted that patients often unintentionally omitted relevant details or assumed they had nothing to report:

> *Some patients say that they do not take any medications, but then later it turns out that they take seven!* [P3]

Some also faced barriers when patients were unwilling or unable to disclose their history due to communication breakdowns:

> *Sometimes there are communication difficulties, particularly with foreign patients who may struggle with the [Swedish] language. Some patients are also unwilling to share their medical history.* [P14]

Proxy Reporting

There are situations where medical information is obtained through proxy reporting, which in this case refers to: direct relatives, companions, or care home personnel. Companions refer to those physically accompanying the patient. There were challenges presented here as well:

> *I remember one patient who arrived with a care home worker who did not know the patient's medical history. The patient had a neurodegenerative condition, so we were unable to obtain the information directly. We tried contacting the care home, but there was no nurse available at the time. Eventually, we received a phone number for a direct relative, who was able to provide us with the necessary information.* [P2]

Medical Consultations

Participants described engaging in medical consultations to obtain relevant information or seek advice on complex cases. Primary care units were the most frequent point of contact, although secondary care was also involved, depending on the patient's condition:

> *I would estimate that I contact primary care units about 1–3 times per month. Occasionally, I call specific doctors directly to confirm details if I suspect*

> *the information provided by the patient is inaccurate or incomplete.* [P4]

> *I contact medical departments several times a month. Hematology for chemotherapy patients, oncology for treatment schedules, as well as nephrology and cardiology for patient-specific matters.* [P9]

However, participants often encountered coordination limitations, such as delayed or absent responses, which hindered decision-making and delayed treatments:

> *The main issue with calling medical professionals is that they are hard to reach. You often have to call multiple times, and that takes time. This often leads to postponing treatments, even for emergency cases. I cannot proceed if I am not sure about a patient's medical history.* [P4]

Some reported being redirected or denied help, further complicating access to essential information:

> *I once called a primary care unit, and they responded strangely. They asked how I had gotten their number. Even after I explained that I was a dental practitioner and provided the necessary patient details, they refused to share any information.* [P5]

To cope with these challenges, some participants began asking patients to coordinate directly with their medical providers:

> *I once tried to contact a pediatrician, but it was very difficult to reach them. They had many patients, and it took nearly a year to get a phone appointment. Since then, I have let the patients handle the contact themselves. […] They [patients] ask their doctor, for example, whether they need antibiotic prophylaxis or if they should stop their blood thinner before an extraction, and then they get back to us with the answer.* [P14]

Current Workflow: Human-Computer Interface

The sociotechnical dimension of the human-computer interface examined the interaction with the EDR system when retrieving, documenting, and managing medical information (Table 4). Overall, participants emphasized the interface's critical role in the efficient and reliable handling of medical information, revealing both challenges and opportunities for improving usability.

Table . Aspects of current interface design and expressed design needs, categorized under the sociotechnical dimension of "human-computer interface."

Themes	Participants
Free text input	
Medical history section	P1, P4, P8, P9, P12, P14, P15, P16
Progress notes	P2, P3
Adaptive documentation behavior	P5, P6, P7, P10, P11, P13, P17, P18
Current interface limitations	
Information reliability	P1, P3, P12, P13, P17, P18
Information visibility	P2, P7, P9, P10, P15, P17
Manual data entry	P4, P6, P11, P12, P13, P14, P16, P17, P18
Design needs	
Annotations	P5, P6, P18
Improved information overview	P6, P15, P16, P17, P18
Confirm review	P6
Cross-disciplinary substance abuse alert	P10, P11
Integrated external data accessibility	P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P14, P15, P16, P17, P18

Free Text Input

All participants relied solely on free text input when documenting medical information in the EDR. This was done either in the medical history section (a structured form with predefined fields) or in progress notes (where other dental care activities are recorded).

> *There is a medical history section form that covers common illnesses like asthma, osteoporosis, cancer, diabetes, cardiovascular disease, liver and kidney diseases, rheumatism, mental health issues, history of radiation therapy, bleeding disorders, allergies, etc. If yes to any question, we ask them what medications they take and type it in.* [P14]

Some participants preferred documenting directly in progress notes due to time constraints and navigational hurdles:

> *In the EDR, we have a medical history section to fill in information. […] However, due to time constraints, I do not have time to go through everything, so I type in the progress notes instead.* [P2]

Most adopted an adaptive documentation behavior, using both sections to reinforce critical information:

> *I use the dedicated medical history section for documenting medical information. If there is something important, like bisphosphonates, I also type it in the progress notes.* [P6]

Current Interface Limitations

The category current interface limitations revealed several usability issues. One major concern was information reliability, where participants questioned the accuracy and completeness of previously entered medical information in the EDR system:

> *The problem with the current system is that we do not know if the medical information is recorded accurately or not by previous colleagues, or even if it is reliable.* [P1]

Another issue was information visibility, where inconsistent documentation across sections made it difficult to locate medical details:

> *If medical information is buried in the progress notes, it takes time to find. Time is also wasted when it is missing or outdated.* [P7]

Manual data entry was also a major limitation. Participants described it as time-consuming and prone to errors, especially when documenting long medication lists:

> *When there is a lot of medical information, entering it manually is a burden because it takes time, which could have been spent with the patient. Also, correctly spelling the names under time pressure can lead to mistakes.* [P17]

Design Needs

The findings in design needs emphasized the need for better system integration, interface enhancements, and improved visibility of patient data. All participants expressed a strong preference for integrated external data accessibility, where medical information should be directly accessible within the EDR:

> *Compared to EHR systems, which display all medications and diagnoses, EDR systems fall short. […] Our work is also time-pressured, which is why an integrated solution providing access to medical information would save time. […] Even if in another program, that is still better than nothing.* [P5]

Some also emphasized the need for access control, ensuring that only authorized personnel could view sensitive data:

> *Ideally, all relevant medical information should be retrieved in a controlled process to safeguard patients' privacy.* [P15]

Additional interface improvements suggested included:

> *Annotations: If there was a built-in description of each entry, that would be incredibly helpful. Like a pop-up suggestion that says what a particular drug is for.* [P5]
>
> *Improved Information Overview: An overview of the [medical] information is preferable, but it must be possible to filter it, so the overview stays clear and focused.* [P17]
>
> *Confirm Review Feature: It would be best if there was a function where you could sign to confirm that you have reviewed the medical history. […] to ensure accountability.* [P6]
>
> *Substance Abuse Alerts: We had a patient […] who seeks emergency care only to obtain narcotic medications. […] In EHR systems, this patient has a warning that flags up for that reason. I would like to see such functionality integrated into the EDR.* [P10]

Current Workflow: Workflow and Communication

Workflow and communication focused on broader clinical activities involving interactions with both the EDR system and patients in the context of managing medical information (Table 5). Overall, the findings indicate that workflow efficiency and patient safety are strongly linked to the availability and clarity of medical information. When such information is incomplete or difficult to retrieve, it disrupts clinical routines and compromises care.

Table . Challenges and limitations of the current workflow, as well as perceived benefits of direct medical information access, categorized under the sociotechnical dimension of "workflow and communication."

Themes	Participants
Current documentation workflow	
Time-consuming	P1, P2, P5, P6, P9, P10, P11, P13, P14, P15, P16, P17, P18
Delay treatments	P1, P2, P3, P4, P5, P6, P8, P11, P13, P16
In-treatment hazards	P1, P4, P7, P8, P13, P18
Patient frustration	P8, P10
Direct information accessibility	
Clinical time management	P1, P2, P4, P5, P7, P8, P9, P10, P11, P12, P13, P14, P16, P17, P18
Improved treatment planning	P3, P7, P8, P17
Cognitive load	P3, P12
Better decision-making	P4, P6, P7, P12, P13, P14, P15, P17, P18

Current Documentation Workflow

Participants overwhelmingly described the current documentation workflow as time-consuming, marked by frequent disruptions when obtaining and recording medical information:

> *Asking patients about their medical history, then reading and documenting it accurately, takes up valuable time from the appointment.* [P18]

These issues often led participants to delay treatments, which often involved canceling or rescheduling appointments:

> *"Sometimes, I cannot proceed with the treatment if I do not have the patient's full medical history. I have to explain to the patient that I cannot do anything at that moment because it would pose a risk to their health.* [P2]

In some instances, essential medical information would emerge during procedures, posing potential in-treatment hazards**:**

> *One time, a patient told me he was healthy, and just before the [tooth] extraction, he casually mentioned […] a serious condition. […] We had to stop the procedure immediately.* [P7]

Finally, repeated appointments caused by information gaps contributed to patient frustration:

> *I had to postpone the treatment […] The coordination turned out to be difficult, and the patient became increasingly frustrated.* [P8]

Direct Information Accessibility

Having direct information accessibility was consistently identified as a key improvement. The most frequently perceived benefit was enhanced clinical time management, as direct access would reduce administrative tasks and allow more focus on patient care:

> *"It would be incredibly helpful to have access to patients' full medical history. Then, I would not need to ask the patients because I would be able to look it up myself directly. That way, I could plan treatments faster instead of calling around for additional information.* [P2]

Others highlighted how improved treatment planning would be possible by enabling better scheduling and avoiding unnecessary visits:

> *If I had direct access, I would be able to plan better. Like, if a patient takes a certain medication, I would know that they need to come within 6 months before their next dose.* [P3]

Also, better decision-making was noted as a distinct benefit. Several participants stressed that it would reduce uncertainty, improve efficiency, and enhance patient safety:

> *Having direct access to medical information would speed up decision-making and let patients get treated faster. I would also feel more secure making the right call.* [P6]

A final point was the reduction of patients' cognitive load, as they would no longer need to recall or bring documentation themselves:

> *Patients would not have to recall or remember to bring documents, which would offload them of responsibilities.* [P3]

Current Workflow: Internal Organizational Policies, Procedures, and Culture

The findings from internal organizational policies, procedures, and culture suggest a need for stronger organizational support, more equitable access to medical information, and support for consistency in information management practices (Table 6). These are in relation to managing medical information within PDSD, but also in the context of Region Dalarna.

Table . Perspectives of organizational and procedural aspects categorized into the sociotechnical dimension of "internal organizational policies, procedures, and culture."

Themes	Participants
Current documentation practices	
Existing routines	P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P12, P13, P14, P15, P16, P17, P18
Inconsistencies	P2, P5, P7, P9, P11, P18
System access limitation	
Inaccessible medication list platform	P3, P4, P5, P6, P8, P11, P12, P13, P14, P16, P17, P18
EHR[a] access inequities	P5, P7, P9, P10, P11, P15, P17
Organizational support	
Organizational inaction	P1, P9, P15
Devalued	P11, P17

[a] EHR: electronic health record.

Current Documentation Practices

In the category of current documentation practices, it was widely acknowledged that there are existing routines in place. The existing routines involved checking and updating information for every appointment, as well as reminding patients to bring their medication lists:

> *We must always complete and review the patient's medical history before starting treatment, in order to check whether anything has changed since their last visit.* [P3]
>
> *The routine is that, when booking an appointment, the receptionist tells patients to bring their medication list.* [P9]

It also involved using a dedicated medical history section of EDR. Yet many participants noted inconsistencies in how these were applied, for example, documenting in progress notes or not updating medical information at all:

> *We have a dedicated medical history section that patients fill out for every visit. Though some colleagues only record information in the progress notes instead of using it. That makes things difficult because you then have to dig through the entire record to find relevant information.* [P5]

System Access Limitations

An initiative mentioned by participants was "Förskrivningskollen" [51], an online platform enabling authorized users to view patients' prescribed medications. However, it was described to be inaccessible due to technical issues or not provide enough information:

> *There is a platform that we technically have access to […]. It is called 'Förskrivningskollen.' While it shows prescriptions, it does not indicate what the patient is actually taking or the reasons behind those prescriptions.* [P4]

The second barrier concerned EHR access inequities. While hospital-based or specialized dental clinics within PDSD have access to EHR systems, GDPs do not. Participants viewed this as an unjustified discrepancy:

> *Dental practitioners working in hospital settings have access to EHR systems […] I asked around why we do not have the same access in GDPs, but no one could give me a clear answer.* [P9]

Organizational Support

Organizational support reflects how participants perceived the level of support provided for managing medical information in dental care, where organizational inaction refers to the lack of follow-through on identified issues, even after concerns were raised, and devalued captures the sense of not being prioritized by the broader organization:

> *People often do not realize that dental treatments also depend on knowing about diseases and medications. I think many decision-makers do not understand our professional needs. We are frequently treated like the poor cousins of medical professionals, but what we do has a big impact. We do not want the information out of curiosity. We need it to work safely and effectively.* [P11]
>
> *Dentists need access to relevant medical information, and physicians also need greater knowledge of how diseases and medications affect oral health. More collaboration is needed, because it is ultimately about the patient.* [P17]

Current Workflow: External Rules, Regulations, and Pressures

The last sociotechnical dimension examined was external rules, regulations, and pressures. What can be derived from Table 7 is a need for clearer guidance and communication regarding applicable regulations in dental care regarding the subject matter.

Table . Overview of the categorization under the sociotechnical dimension of "external rules, regulations, and pressures."

Themes	Participants
External regulations	
Awareness	P1, P2, P3, P5, P6, P10, P11, P14, P16
Standard practice	P4, P7, P8, P9, P12, P15, P17, P18
Patient consent	P5, P6, P11, P12, P14, P17, P18

External Regulations

The results show that there is uncertainty about external regulations guiding the management of medical information in dental care. While many participants demonstrated awareness of existing regulations, several assumed that obtaining and documenting medical information is standard practice:

> *I do not know of any rules that tell us explicitly to record medical history. I think it is obvious that we need to.* [P9]
>
> *National guidelines require us to document patients' medical information.* [P16]

Furthermore, the role of patient consent was perceived as crucial for retrieving medical information from both internal and external sources:

> *Yes, the Patient Data Act states that dental practitioners and hygienists may access medical information, but only what is necessary for [providing] dental care. However, if we want to know patients' medical history, we need their consent.* [P14]

Current Workflow: Conceptual Model

A UML activity diagram was developed to illustrate the current workflow for managing medical information in dental care based on the interview data (Multimedia Appendix 3). It outlines the key steps involved in acquiring, updating, and documenting medical information in the EDR system before proceeding with treatments. The diagram includes multiple actors, including receptionist, patient, dental practitioner, accompanying proxy, proxy, and medical care unit, each with defined responsibilities represented within separate partitions.

The workflow begins with the receptionist scheduling the appointment and reminding the patient to bring their medication list. Upon arrival, the dental practitioner checks the EDR to determine whether the patient is new or returning. For returning patients, the medical history section and progress notes are reviewed to assess whether updates are needed. If updates are required, or the patient is new, the dental practitioner checks if an accompanied proxy is present and able to provide the necessary information. If not, alternative steps may involve contacting another proxy or a relevant Medical Care Unit.

If no proxy is involved, the dental practitioner evaluates whether the patient can provide sufficient information via verbal recall, a medication list, or their online patient portal. If these sources are insufficient, attempts are made to contact a proxy or medical care unit, with the option to postpone the appointment.

Once the necessary information is obtained, it is documented in the EDR, either in the medical history section or progress notes. The dental practitioner may seek further medical consultation if needed. Once all relevant information is confirmed and documented, treatment can proceed.

Potential Workflow: Conceptual Model

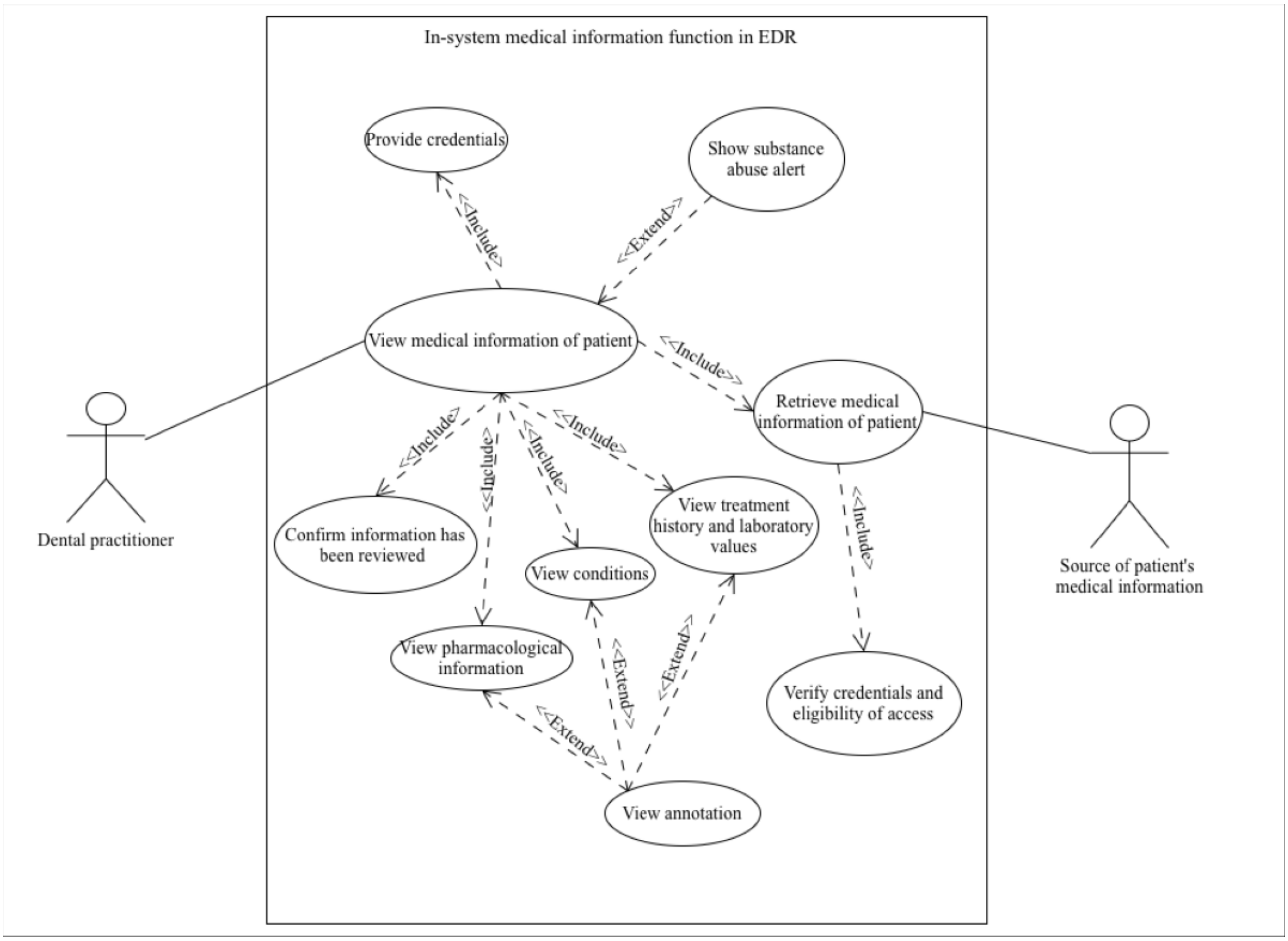
A UML activity diagram of a proposed workflow reflects potential improvements if medical information were made accessible in the EDR system (Multimedia Appendix 4). This revised workflow removes several steps previously required to manage information gaps and instead introduces a more linear progression toward care provision.

The dental practitioner begins by accessing and checking the medical information of the patient in their EDR system, with the receptionist not having to remind the patient to bring their medication list. Furthermore, there is no reliance on an accompanied proxy, proxy, and medical care unit to obtain medical information. Only if further clarification or consultancy is needed does the process escalate to contacting the medical care unit.

Use Case Related to the Potential Workflow

A UML use case diagram (Figure 1) was developed to represent conceptual expectations and functional needs described by participants regarding an EDR system with access to medical information. This model also assumes that medical information is accessible via a regional database or NPÖ through NTjP, referred to as the source of the patient’s medical information in the diagram. It is important to emphasize that the use case diagram does not depict a user interface. Rather, it consolidates qualitative insights into conceptual recommendations. System-level specifications and exceptions are therefore not included.

Figure 1. A use case with an electronic dental record system that involves providing access to medical information. EDR: electronic dental record.



Below follows the use case, with its corresponding process flow detailed in Table 8**:**

1. Use case name to view the medical information of a patient.
2. Description to describe the steps required for a dental practitioner to securely retrieve and review a patient's medical information via an in-system function in the EDR system.
3. Alternative flow to ensure controlled data access, step 5 includes an alternative flow where, if the forwarded credentials are invalid or belong to a user without authorization, the data request is not processed.
4. Preconditions: the patient has given consent for a dental practitioner to access and view information, and the dental practitioner is directly involved in their care.
5. Postconditions: the patient's medical information has been securely accessed and reviewed by the dental practitioner. The system has logged the confirmation of the reviewed information.

Table . Normal flow of events for accessing medical information.

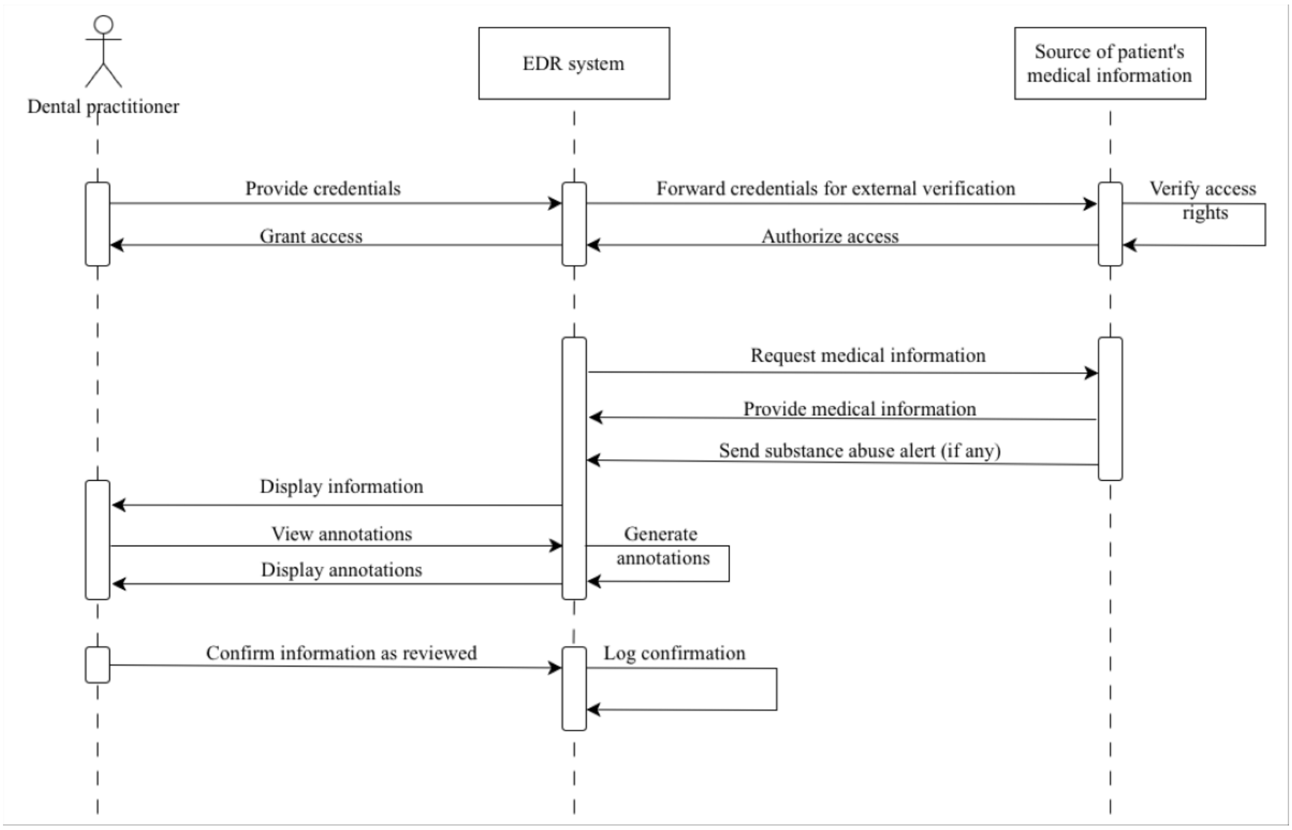
Steps	Actors	Actions
1	Dental practitioner	Initiates action to view the patient's medical information using an in-built function within the EDR[a] system
2	In-system function	Prompts user authentication to verify access rights
3	Dental practitioner	Provides credentials
4	In-system function	Sends a data access request to the external system, including credentials
5	Source of patient's medical information	Validates the access request and verifies that the credentials are valid and authorized to access medical information
6	Source of patient's medical information	Retrieves and returns relevant medical data
7	In-system function	Displays relevant data in the conditions, pharmacological information, and treatment history and laboratory values sections
8	Dental practitioner	Reviews the provided information
9	Dental practitioner	Confirms that the information has been reviewed
10	In-system function	Logs the confirmation
11	Not applicable	End of use case

[a] EDR: electronic dental record.

Sequential Flow of Accessing Medical Information in Potential Workflow

As shown in Figure 2, the UML sequence diagram depicts the interaction between the dental practitioner, the EDR system, and the source of patient's medical information. After the dental practitioner provides credentials to the EDR system, a verification request is sent to the external source. Upon successful authorization, the EDR system sends a request and receives the patient's medical information as a message. This may include substance abuse alerts, if applicable. The EDR system then displays the retrieved data to the dental practitioner, who may also choose to view system-generated annotations. Finally, the dental practitioner confirms that the information has been reviewed, which is logged in the system.

Figure 2. The sequence of processes in facilitating medical information. EDR: electronic dental record.



Discussion

Principal Results

Identified Medical Information Needs

Overall, the interviews consistently highlighted the need for accessible medical information to support safe, timely, and informed dental care provision. The findings predominantly aligned with what is reported in prior literature [3-5]. However, this study provided additional granularity regarding the specific medical information needs and their clinical relevance. Participants emphasized 4 main categories: conditions, pharmacological information, treatment history, and laboratory values.

Medical conditions such as cardiovascular disease, cancer, and psychiatric disorders were repeatedly cited due to their implications for procedural risk and treatment planning. Pharmacological needs centered on antiresorptive agents and anticoagulants, given their link to complications such as bleeding and osteonecrosis. Substance abuse history was also flagged as critical, particularly in the context of narcotic misuse. Allergies were a concern, especially penicillin allergy, due to its prescribing relevance. Moreover, relevant medical treatment history and laboratory values were also considered essential for planning and safety.

Identified Workflow Challenges

At GDPs in Dalarna, the current workflows are hindered by several challenges. These include a reliance on patients or third parties, which often leads to incomplete information gathering. Additionally, documentation workflows are fragmented, characterized by frequent interruptions and manual redundancy. The use of EDR for documenting medical information is inconsistent, and there is limited access to supportive information systems. These challenges represent a sociotechnical gap, where systems exist to store medical information but are not designed to meet the information needs or workflow constraints of the participants. In other words, the mere availability of medical information in EDR systems is insufficient, as usability, reliability, and contextual alignment were regarded as critical elements.

Workflow Barriers and Limitations

The findings underscore a misalignment between the current system design, organizational priorities, and clinical expectancy. Participants described relying heavily on patient self-reports, proxy input, or contacting external medical care units to obtain information. Patient self-reports were often unreliable, especially with elderly patients. Proxy or medical consultations introduced additional delays and communication gaps.

Despite the existence of routines, widespread inconsistencies in how medical histories are recorded were reported. Participants believed that having direct access to relevant medical information via the EDR would enhance treatment planning and reduce delays. However, they felt unsupported by the organization**,** citing failed efforts such as the online prescription check platform “Förskrivningskollen.”

The EDR system itself relies heavily on manual input, causing duplicated entries between the medical history section and progress notes. Manual data entry was perceived as burdensome, error-prone, and inefficient. Inconsistent documentation practices also impaired information reliability and visibility, requiring participants to spend additional time navigating and interpreting records.

There was also uncertainty about legal permissions for accessing medical records, despite the existence of the Integrated Health and Social Care Documentation Act (2022:913) and explicit clarifications from Swedish authorities [21,52]. This points to organizational and educational shortcomings rather than legal barriers. The forthcoming European health data space regulation may help address such ambiguity at the European Union level [53].

Proposed Improvements for Information Access

Enabling direct access to medical information via the EDR could streamline workflows by reducing reliance on manual data retrieval. This would improve time management, decision-making, and care coordination. A feasible path forward involves leveraging Sweden’s national interoperability infrastructure. By connecting EDR systems to NPÖ through NTjP, authorized dental professionals could retrieve relevant medical data across institutional boundaries [22-24].

Addressing Interoperability Levels

To implement this vision, interoperability must be achieved at four levels [11]: (1) technical, which must be supported through NTjP; (2) syntactic, which requires EDR systems to align with NTjP service contracts; (3) semantic, to provide a view-only use of external medical information, due to current dental records being largely unstructured; and (4) organizational, which requires commitment from vendors and stakeholders to align with interoperability frameworks and recognize dental professionals’ medical information access needs.

Toward a User-Centered Design Approach

Interface inefficiencies and inconsistent documentation stem not only from system limitations but also from design misalignment. Poorly designed systems, even with access to the right data, can obstruct rather than support care [3,13]. Given the high cognitive load and time pressure in dental care [54], systems should minimize user burden. Preferences expressed by participants underscore the need for user-centered EDR design regarding medical information.

Implications for Scalability and Policy Considerations

In highly urbanized or well-resourced Swedish regions, the proposed integration model would require significant adaptations to ensure scalability. These environments typically involve higher patient volumes, greater clinical complexity, and a more diverse ecosystem of EHR and EDR systems. To operate effectively under such conditions, the model must support real-time data synchronization, load balancing, and robust multisystem interoperability through standardized APIs, for example, applying the Health Level Seven Fast Healthcare Interoperability Resources standard and a modular system architecture. Larger regions may also necessitate granular, role-based access controls and comprehensive audit trails to uphold security and regulatory compliance. Scalable infrastructure, potentially leveraging cloud-based or hybrid deployment models, will be essential to accommodate periods with high system workload.

Beyond technical enhancements, successful large-scale implementation will depend on policy and infrastructural measures. These include enforcing national interoperability and terminology standards (eg, Fast Healthcare Interoperability Resources and Systematized Nomenclature of Medicine Clinical Terms), updating the Swedish Patient Data Act to explicitly enable cross-sector data sharing between dental and medical care while maintaining General Data Protection Regulation compliance, and allocating targeted funding to modernize legacy dental systems and support workforce training. Together, these considerations enhance the workflow’s scalability and support its potential application in settings that differ from the region studied.

Challenges of Integrating Artificial Intelligence–Driven Dental Imaging and Analysis Into EDR

Beyond the information needs described in this study, the rapidly expanding use of specialized dental software, such as artificial intelligence (AI)-driven digital smile design systems [55] and video-based smile analysis tools like Dynasmile [56], introduces additional challenges for integrating diverse data types into EDRs. These applications generate complex non-textual outputs, including high-resolution images, videos, and AI-derived esthetic or diagnostic metrics, which are increasingly used for treatment planning. Recent reviews of AI-based digital smile design indicate promising improvements in patient and clinician satisfaction, as well as esthetic outcomes [57]. Integrating such data into EDRs requires overcoming significant interoperability barriers. For instance, digital imaging data must comply with standards like Digital Imaging and Communications in Medicine to ensure consistent storage, retrieval, and portability across systems [58]. Similarly, AI-derived outputs (eg, landmark coordinates and esthetic scores) need semantic mapping to standardized clinical concepts to ensure meaningful linkage with diagnostic and treatment data. Furthermore, exchanging large multimedia files through NTjP raises practical considerations regarding storage, capacity, bandwidth, consent, and privacy compliance [59].

Acknowledging these challenges underscores the broader implications of our findings and highlights the need for future research on incorporating advanced dental imaging and AI-generated data into EDR workflows without compromising usability, interoperability, or patient confidentiality.

Limitations

Certain limitations need to be declared. First, the sample consisted of 18 dental practitioners from PDSD-operated GDPs in Dalarna. Although the generalizability was limited due to the limited number of participants and the homogeneity of the sample, with most participants having fewer than five years of professional experience and working in similar clinical settings, thematic saturation was achieved in this study. Qualitative studies, besides providing an in-depth understanding of contextualized experience, may also support theoretical generalizability or, more appropriately for qualitative research, transferability when rich contextual information is provided [60]. This study supports the transferability of the findings to other contexts with similar characteristics. However, to strengthen the generalizability of these findings, future studies should include a broader and more diverse sample. A quantitative or mixed-methods study would also be valuable to validate and generalize the medical information needs identified here [61].

Future research could use quantitative methods to validate and prioritize the information needs identified in this study across a broader population of dental practitioners. Large-scale surveys using structured instruments (eg, Likert-scale items) could assess both the perceived importance and frequency of specific information needs, providing estimates of their prevalence and variability at regional or national levels. Additionally, a Delphi approach incorporating iterative rounds of quantitative scoring could facilitate structured expert consensus and systematic ranking of priorities.

Conclusions

This study identified a clear need to improve access to medical information within EDR systems in GDPs in Dalarna. While medical information needs such as diagnoses, medications, and allergies are widely acknowledged in the literature, this study contributed detailed insights into the types of information most relevant to dental care. Current documentation practices were found to be fragmented, inefficient, and overly reliant on patients and third parties. Usability limitations in the EDR interface further hinder information access and consistency. Collectively, these issues reflect a sociotechnical misalignment between user needs and system capabilities. To address these gaps, the study proposes leveraging Sweden’s existing interoperability infrastructure, particularly by integrating EDR systems with NPÖ via NTjP. However, achieving this requires not only technical alignment but also organizational commitment, regulatory clarity, and a user-centered system design.

Future research should validate these findings at scale and explore system design and policy solutions to ensure that medical information can be securely, efficiently, and meaningfully integrated into information systems used in dental care.

Acknowledgments

The authors want to thank the Public Dental Service Dalarna (PDSD) for their cooperation, and to all the participants who took the time to contribute to this study. It needs to be declared that the main author is a licensed dental practitioner employed by PDSD. However, employment at PDSD did not affect the research process, as no funding was received. Interactions with PDSD were solely limited to recruitment coordination.

Funding

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

Authors' Contributions

SHR conducted the data collection and carried out the initial analysis, which was subsequently reviewed and refined in collaboration with ND and SB.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[PDF File, 124 KB - humanfactors_v13i1e82691_app1.pdf]

Multimedia Appendix 2
Consent form in Swedish and English.
[PDF File, 116 KB - humanfactors_v13i1e82691_app2.pdf]

Multimedia Appendix 3
Unified modeling language activity diagram of the current workflow in managing medical information.
[PNG File, 666 KB - humanfactors_v13i1e82691_app3.png]

Multimedia Appendix 4
Unified modeling language activity of an improved workflow if medical information were made accessible directly in the electronic dental record.
[PNG File, 229 KB - humanfactors_v13i1e82691_app4.png]

References

1. Schmalz G, Brauer L, Haak R, Ziebolz D. Evaluation of a concept to classify anamnesis-related risk of complications and oral diseases in patients attending the clinical course in dental education. BMC Oral Health 2023 Aug 29;23(1):609. [doi: 10.1186/s12903-023-03343-x] [Medline: 37644402]
2. Harrel SK, Molinari J. Aerosols and splatter in dentistry: a brief review of the literature and infection control implications. J Am Dent Assoc 2004 Apr;135(4):429-437. [doi: 10.14219/jada.archive.2004.0207] [Medline: 15127864]
3. Li S, Felix Gomez GG, Xu H, Rajapuri AS, Dixon BE, Thyvalikakath T. Dentists' information needs and opinions on accessing patient information via health information exchange: survey study. JMIR Form Res 2024 Jan 11;8:e51200. [doi: 10.2196/51200] [Medline: 38206667]
4. Simon L, Obadan-Udoh E, Yansane AI, et al. Improving oral-systemic healthcare through the interoperability of electronic medical and dental records: an exploratory study. Appl Clin Inform 2019 May;10(3):367-376. [doi: 10.1055/s-0039-1688832] [Medline: 31141831]
5. Alanazi A, Alghamdi G, Aldosari B. Informational needs for dental-oriented electronic health records from dentists' perspectives. Healthcare (Basel) 2023 Jan 14;11(2):266. [doi: 10.3390/healthcare11020266] [Medline: 36673634]
6. Lisiecka MZ. Allergic reactions in dental practice: classification of medicines, mechanisms of action, and clinical manifestations. Clin Rev Allergy Immunol 2025 Feb 15;68(1):17. [doi: 10.1007/s12016-025-09032-7] [Medline: 39954171]
7. Adibi S, Li M, Salazar N, et al. Medical and dental electronic health record reporting discrepancies in integrated patient care. JDR Clin Trans Res 2020 Jul;5(3):278-283. [doi: 10.1177/2380084419879387] [Medline: 31560579]
8. Charangowda BK. Dental records: an overview. J Forensic Dent Sci 2010 Jan;2(1):5-10. [doi: 10.4103/0974-2948.71050] [Medline: 21189983]
9. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2016 May 20;Suppl 1(Suppl 1):S48-S61. [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
10. Heid DW, Chasteen J, Forrey AW. The electronic oral health record. J Contemp Dent Pract 2002 Feb 15;3(1):43-54. [Medline: 12167912]
11. Lehne M, Sass J, Essenwanger A, Schepers J, Thun S. Why digital medicine depends on interoperability. NPJ Digit Med 2019;2:79. [doi: 10.1038/s41746-019-0158-1] [Medline: 31453374]
12. Holmgren AJ, Adler-Milstein J. Health information exchange in us hospitals: the current landscape and a path to improved information sharing. J Hosp Med 2017 Mar;12(3):193-198. [doi: 10.12788/jhm.2704] [Medline: 28272599]
13. Taylor HL, Apathy NC, Vest JR. Health information exchange use during dental visits. AMIA Annu Symp Proc 2020;2020:1210-1219. [Medline: 33936497]
14. Li S, Rajapuri AS, Felix Gomez GG, Schleyer T, Mendonca EA, Thyvalikakath TP. How do dental clinicians obtain up-to-date patient medical histories? modeling strengths, drawbacks, and proposals for improvements. Front Digit Health 2022;4:847080. [doi: 10.3389/fdgth.2022.847080] [Medline: 35419556]
15. Atchison KA, Weintraub JA, Rozier RG. Bridging the dental-medical divide: case studies integrating oral health care and primary health care. J Am Dent Assoc 2018 Oct;149(10):850-858. [doi: 10.1016/j.adaj.2018.05.030] [Medline: 30057150]
16. About the Swedish healthcare system. Socialstyrelsen. 2019. URL: https://www.socialstyrelsen.se/en/about-us/healthcare-for-visitors-to-sweden/about-the-swedish-healthcare-system/ [accessed 2025-02-01]
17. Uppföljning Vision e-hälsa 2025 – Rapport avseende år 2021 (Report in Swedish). : E-hälsomyndigheten; 2022 URL: https://www.ehalsomyndigheten.se/globalassets/ehm/3_om-oss/rapporter/uppfoljning-vision-e-halsa-2025-rapport-avseende-2021.pdf [accessed 2026-01-27]
18. Follow-Up Vision for eHealth 2025 - Report on the Year 2022. : Swedish eHealth Agency; 2023 URL: https://www.ehalsomyndigheten.se/siteassets/ehm/3_om-ehalsomyndigheten/rapporter/rapporter_regeringsuppdrag/follow-up-vision-for-ehealth-2025-report-on-the-year-2022.pdf [accessed 2026-01-27]
19. Swedish Institute. Swedish government. Power for the people! This is how Sweden is governed. URL: https://sweden.se/life/democracy/swedish-government [accessed 2025-02-12]
20. Nilsson P, Stjernquist A, Janlöv N. Fragmented health and social care in Sweden - a theoretical framework that describes the disparate needs for coordination for different patient and user groups. Int J Integr Care 2016;16(6):348. [doi: 10.5334/ijic.2896]
21. The Government Offices' legal databases [Web page in Swedish]. Regeringskansliet. 2022. URL: http://rkrattsbaser.gov.se/sfst?bet=2022:913 [accessed 2025-02-14]
22. Sammanhållen journalföring – Möjligheter till digital informationsförsörjning på hälsodataområdet. : E-hälsomyndigheten; 2022 URL: https://www.ehalsomyndigheten.se/globalassets/ehm/3_om-oss/rapporter/slutrapportering-av-uppdraget-att-foresla-hur-sammanhallen-journalforing-kan-nyttjas-i-storre-utstrackning.pdf [accessed 2025-02-14]
23. Nationell patientöversikt – NPö [Web page in Swedish]. Inera. URL: https://www.inera.se/tjanster/alla-tjanster-a-o/npo---nationell-patientoversikt/ [accessed 2025-02-14]
24. Nationella tjänsteplattformen. Inera. URL: https://www.inera.se/tjanster/arkitektur-och-digital-infrastruktur/nationella-tjansteplattformen/ [accessed 2025-02-14]
25. Långt kvar till gemensamt IT-system [Web page in Swedish]. Tandläkartidningen. 2017. URL: https://www.tandlakartidningen.se/nyhet/langt-kvar-till-gemensamt-it-system/ [accessed 2025-03-22]

26. Atchison KA, Weintraub JA. Integrating oral health and primary care in the changing health care landscape. N C Med J 2017;78(6):406-409. [doi: 10.18043/ncm.78.6.406] [Medline: 29203606]
27. MacNeil RLM, Hilario H, Ryan MM, Glurich I, Nycz GR, Acharya A. The case for integrated oral and primary medical health care delivery: marshfield clinic health system. J Dent Educ 2020 Aug;84(8):924-931. [doi: 10.1002/jdd.12289] [Medline: 32638376]
28. Shimpi N, Glurich I, Panny A, Acharya A. Knowledgeability, attitude, and practice behaviors of primary care providers toward managing patients' oral health care in medical practice: Wisconsin statewide survey. J Am Dent Assoc 2019 Oct;150(10):863-872. [doi: 10.1016/j.adaj.2019.05.020] [Medline: 31446976]
29. Harnagea H, Couturier Y, Shrivastava R, et al. Barriers and facilitators in the integration of oral health into primary care: a scoping review. BMJ Open 2017 Sep 25;7(9):e016078. [doi: 10.1136/bmjopen-2017-016078] [Medline: 28951405]
30. Unika siffror: här är tandläkarbristen värst. Tandläkartidningen. 2023. URL: https://www.tandlakartidningen.se/nyhet/unika-siffror-har-ar-tandlakarbristen-varst/ [accessed 2025-02-22]
31. Dalarna. Tillväxtverket. URL: https://tillvaxtverket.se/tillvaxtverket/statistikochanalys/statistikomregionalutveckling/lansuppdeladstatistik/dalarna.1593.html?chartCollection=6#svid12_843740f184cc4e52fa14d4 [accessed 2025-02-22]
32. Janto M, Iurcov R, Daina CM, et al. Oral health among elderly, impact on life quality, access of elderly patients to oral health services and methods to improve oral health: a narrative review. J Pers Med 2022 Feb 28;12(3):372. [doi: 10.3390/jpm12030372] [Medline: 35330372]
33. Alami H, Lehoux P, Gagnon MP, Fortin JP, Fleet R, Ag Ahmed MA. Rethinking the electronic health record through the quadruple aim: time to align its value with the health system. BMC Med Inform Decis Mak 2020 Feb 17;20(1):32. [doi: 10.1186/s12911-020-1048-9] [Medline: 32066432]
34. Baxter G, Sommerville I. Socio-technical systems: from design methods to systems engineering. Interact Comput 2011 Jan;23(1):4-17. [doi: 10.1016/j.intcom.2010.07.003]
35. Hunter DJ, McCallum J, Howes D. Defining exploratory-descriptive qualitative (EDQ) research and considering its application to healthcare. Proceedings of the 6th Annual Worldwide Nursing Conference 2018 2018;4(1) [FREE Full text]
36. Tandvård genom hela livet [Web page in Swedish]. Region Dalarna. URL: https://www.regiondalarna.se/verksamhet/tandvard/ [accessed 2025-03-22]
37. Ordell S, Unell L, Söderfeldt B. An analysis of present dental professions in Sweden. Swed Dent J 2006;30(4):155-164. [Medline: 17243443]
38. 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. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
39. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019;7(2):e000057. [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
40. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Qual Saf Health Care 2010 Oct;19 Suppl 3(Suppl 3):i68-i74. [doi: 10.1136/qshc.2010.042085] [Medline: 20959322]
41. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
42. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
43. Neale J. Iterative categorization (IC): a systematic technique for analysing qualitative data. Addiction 2016 Jun;111(6):1096-1106. [doi: 10.1111/add.13314] [Medline: 26806155]
44. Ciccozzi F, Malavolta I, Selic B. Execution of UML models: a systematic review of research and practice. Softw Syst Model 2019 Jun;18(3):2313-2360. [doi: 10.1007/s10270-018-0675-4]
45. Aggarwal V. The application of the unified modeling language in object-oriented analysis of healthcare information systems. J Med Syst 2002 Oct;26(5):383-397. [doi: 10.1023/a:1016449031753] [Medline: 12182204]
46. Lag om etikprövning av forskning som avser människor [Web page in Swedish]. SVERIGES RIKSDAG. URL: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/lag-2003460-om-etikprovning-av-forskning-som_sfs-2003-460/ [accessed 2025-08-05]
47. What the Act says [Web page in Swedish]. Etikprövningsmyndigheten. URL: https://etikprovningsmyndigheten.se/en/what-the-act-says/ [accessed 2025-02-14]
48. European Parliament and Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Off J Eur Union 2016;L119:1-88 [FREE Full text]
49. Association WM. 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]
50. Om oss [Web page in Swedish]. Carestream dental. URL: https://www.t4academy.se/pages/om-oss [accessed 2025-04-05]
51. Förskrivningskollen [Web page in Swedish]. E-hälsomyndigheten. URL: https://www.ehalsomyndigheten.se/yrkesverksam/vard/forskrivningskollen/ [accessed 2025-03-20]

52. Kartläggning av hinder för samverkan mellan tandvård och hälso- och sjukvård – slutrapport 2019 [Web page in Swedish]. Socialstyrelsen. 2019. URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2019-10-6428.pdf [accessed 2025-02-14]
53. European Health Data Space Regulation (EHDS). European Commission. URL: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en [accessed 2025-05-13]
54. Pontes CC, Stanley K, Molayem S. Understanding the dental profession's stress burden: prevalence and implications. Compend Contin Educ Dent 2024 May;45(5):236-241. [Medline: 38900461]
55. Khaled Ali GA, Abdelbadei KA, Essam M, Ali GAK, et al. Smile.AI: a deep learning system for digital smile design. : IEEE Presented at: 2023 Eleventh International Conference on Intelligent Computing and Information Systems (ICICIS); Nov 21-23, 2023; Cairo, Egypt p. 568-573. [doi: 10.1109/ICICIS58388.2023.10391194]
56. Chen K, Qiu L, Xie X, Bai Y. Dynasmile: video-based smile analysis software in orthodontics. SoftwareX 2025 Feb;29:102004. [doi: 10.1016/j.softx.2024.102004] [Medline: 40438663]
57. Saini RS, Kaur K, Gurumurthy V, et al. Impact of artificial intelligence-based digital smile design on patient and clinician satisfaction and facial esthetic outcomes: a systematic review and meta-analysis. Digit Health 2025;11. [doi: 10.1177/20552076251388392] [Medline: 41122427]
58. Farman AG. Use and implication of the DICOM standard in dentistry. Dent Clin North Am 2002 Jul;46(3):565-573. [doi: 10.1016/s0011-8532(02)00003-4] [Medline: 12222098]
59. Liu TY, Lee KH, Mukundan A, Karmakar R, Dhiman H, Wang HC. AI in dentistry: innovations, ethical considerations, and integration barriers. Bioengineering (Basel) 2025 Aug 29;12(9):41007172. [doi: 10.3390/bioengineering12090928] [Medline: 41007172]
60. Carminati L. Generalizability in qualitative research: a tale of two traditions. Qual Health Res 2018 Nov;28(13):2094-2101. [doi: 10.1177/1049732318788379] [Medline: 30043686]
61. Polit DF, Beck CT. Generalization in quantitative and qualitative research: myths and strategies. Int J Nurs Stud 2010 Nov;47(11):1451-1458. [doi: 10.1016/j.ijnurstu.2010.06.004] [Medline: 20598692]

Abbreviations

AI: artificial intelligence
EDR: electronic dental record
EHR: electronic health record
GDP: general dental practice
HIE: health information exchange
IEHR: integrated EHR-EDR
NPÖ: nationell patientöversikt
NTjP: nationell tjänsteplattform
PDSD: Public Dental Service Dalarna
UML: unified modeling language

Edited by A Kushniruk; submitted 20.Aug.2025; peer-reviewed by K Chen, S Li; revised version received 05.Dec.2025; accepted 14.Dec.2025; published 29.Jan.2026.

Please cite as:
Rahim SH, Davoody N, Bonacina S
Exploring Medical Information Needs and Accessibility in Swedish Dental Care by Analysis of Documentation Workflows and Electronic Dental Records in Dalarna: Sociotechnical Qualitative Study
JMIR Hum Factors 2026;13:e82691
URL: https://humanfactors.jmir.org/2026/1/e82691
doi:10.2196/82691

Evaluating User Engagement and Satisfaction With Digital Mental Health Interventions: Randomized Controlled Trial of a Text Messaging Program and e-Mental Health Resources

Gloria Obuobi-Donkor[1], PhD; Reham Shalaby[2], PhD, MD; Belinda Agyapong[2], MEd, PhD; Samuel Obeng Nkrumah[1], MD, MSc; Medard Kofi Adu[1], PhD; Ejemai Eboreime[1], PhD, MD; Lori Wozney[1,3], PhD; Vincent Israel Opoku Agyapong[1,2], PhD, MD

[1]Department of Psychiatry, Faculty of Medicine, Dalhousie University, 5909 Veterans Memorial Lane, 8th Floor. Abbie J. Lane Memorial Building QEII Health Sciences Centre, Halifax, NS, Canada

[2]Department of Psychiatry, Faculty of Medicine & Dentistry, University of Alberta, Edmonton, AB, Canada

[3]Mental Health and Addictions Program, IWK Health, Halifax, NS, Canada

Corresponding Author:
Vincent Israel Opoku Agyapong, PhD, MD
Department of Psychiatry, Faculty of Medicine, Dalhousie University, 5909 Veterans Memorial Lane, 8th Floor. Abbie J. Lane Memorial Building QEII Health Sciences Centre, Halifax, NS, Canada

Abstract

Background: Digital mental health tools, such as SMS text messaging and online resources, are increasingly used to support well-being. However, user satisfaction across these formats remains insufficiently explored.

Objective: The study assessed participants' engagement, perceived impact, and overall satisfaction with the Text4Support program and the e-mental health resources.

Methods: This randomized controlled study was conducted in Nova Scotia, Canada. Participants were assigned to either the Text4Support group, which received daily supportive text messages, or the control group, which received a single text message with a link to the Nova Scotia Mental Health and Addiction Program e-mental health resources. Responses to various aspects of the interventions were evaluated using a 5-point Likert scale, while overall satisfaction was measured on a scale from 0 to 10. The chi-square test and Fisher exact test were employed for data analysis.

Results: A total of 69 participants in the control group and 130 in the Text4Support group completed the satisfaction survey. The overall mean (SD) satisfaction score in the control group was 5.1 (2.3), and the overall mean (SD) satisfaction score for the Text4Support group was 7.1 (2.2). Compared to the control group, participants in the Text4Support group reported greater engagement and positive program impact. While 53.8% (70/130) of Text4Support recipients always read the messages, only 39.1% (27/69) of the control group rarely accessed the eHealth resources. When compared to the control group, participants allocated to the Text4Support group were reported to sometimes take positive action upon reading the messages (42.3% vs 33.3%). A significantly higher proportion of Text4Support users strongly agreed or agreed that the messages were supportive (81.4% vs 41.5%), positive (88.4% vs 49.2%), and helpful in coping with stress (44.2% vs 11.9%), loneliness (40.3% vs 13.4%), and improving mental well-being (51.2% vs 17.9%). In contrast, the majority of responses from the control group were largely neutral.

Conclusions: Results showed that Text4Support group participants were significantly more satisfied with the program than those receiving standard eHealth resources. This highlights that daily supportive SMS text messaging is an effective, low-cost adjunct to care delivery and mental health improvement. These findings suggest that aggregate, brief, and low-cost text-based interventions have great potential for increasing health access and engagement, particularly among traditionally disadvantaged populations with limited access to traditional services.

Trial Registration: ClinicalTrials.gov NCT05411302; https://clinicaltrials.gov/study/NCT05411302

(JMIR Hum Factors 2026;13:e76587) doi:10.2196/76587

KEYWORDS

text messaging; online resources; mental health; Nova Scotia; user satisfaction

Introduction

Psychiatric care has increasingly incorporated digital mental health interventions, which offer accessible and scalable support beyond traditional therapy and medication. Patient feedback through user satisfaction measures is vital in assessing the effectiveness of digital mental health interventions [1]. Patient experiences are gaining recognition as a key factor in enhancing health care quality alongside clinical effectiveness and safety [2,3]. These patient care experiences are included in public reporting and health performance programs [3]. Supportive SMS text messaging programs and eHealth services have become popular and convenient tools for managing mental health [4,5]. The structure and approaches of these interventions may vary in terms of user engagement and perceived effectiveness, but all aim to enhance well-being [6]. Nevertheless, satisfaction levels often vary, highlighting an area that requires more attention to optimize the effectiveness of such interventions for individuals with mental illness.

Supportive SMS text messaging programs like Text4Hope, Text4Mood, Text4PTSI, and Wellbeing offer automated text messages that provide preprogrammed positive reinforcement to users [7-9]. These coping message text programs require minimal input from users while guiding users through daily coping skills and emotional support [7-11]. Literature has shown that these programs increase feelings of connectedness and reduce mental health conditions [10]. Two randomized controlled trials (RCTs) conducted in Canada revealed that supportive text messages reduced depression and alcohol intake of subscribers [12,13].

Various studies have shown that eHealth resources such as online self-help services, psychoeducational tools, and virtual peer support groups provide an accessible way to seek help, as well as an independent role in receiving care [14,15]. These entail being able to partake in the content through specific timestamps, which ensures self-governance [16]. Nevertheless, there is little standardization regarding how users engage with eHealth interventions; it remains their prerogative to seek and employ the information furnished, and this influences overall satisfaction and perceived efficacy.

. This study aims to compare supportive SMS text messaging interventions with eHealth resources to better understand users' perceptions and feedback on the impact, engagement levels, and overall satisfaction. The study will evaluate engagement and user satisfaction with Text4Support in comparison to eHealth resources and assess the perceived relevance and impact of supportive SMS text messaging relative to these eHealth tools.

Given the importance of the perceived adequacy of digitally delivered tools as a core component of modern psychiatric intervention delivery systems to their success, it is essential to understand how users evaluate the adequacy of digitally delivered interventions. This study aims to compare supportive SMS text messaging interventions with eHealth resources to better understand users' perceptions and feedback on the impact, engagement levels, and overall satisfaction. The study will evaluate engagement and user satisfaction with Text4Support in comparison to eHealth resources and assess the perceived relevance and impact of supportive SMS text messaging relative to these eHealth tools.

The intervention, Text4Support, provides daily cognitive behavioral therapy (CBT)-informed supportive text messages for 6 months, tailored to users' primary mental health concerns and delivered as an adjunct to usual care. In contrast, the control group receives the usual care plus a single text message directing participants to freely available e-mental health tools offered through the Nova Scotia Mental Health and Addictions Program website.

Methods

Study Design

This RCT adopted a multicenter, two-arm design with a rater-blinded methodology. Participants were recruited and randomized into the Text4Support group and the control group.

Ethical Considerations

The study was approved by the Nova Scotia Health Research Ethics Board (REB File #1028174) and registered with ClinicalTrials.gov (NCT05411302). Each study participant provided informed consent, and the study was conducted according to the Declaration of Helsinki [17] and Good Clinical Practice (Canadian Guidelines) [18]. Participation was voluntary, with the option to withdraw at any time without impact on care. Participants' privacy and confidentiality were rigorously protected. All data were deidentified at the point of collection, stored, and secured. No compensation or incentives were provided to participants.

Data Collection/Intervention

The published study protocol detailed the data collection procedure [19]. In summary, data were collected via a self-administered survey powered by the Research Electronic Data Capture (REDCap) software program [20]. Participants' age, gender, income, education, housing status, relationship status, and clinical information were obtained. Data were collected between October 11, 2022 and July 31, 2024. After thoroughly informing participants about the study, we used an electronic consent form to obtain their consent.

Participants randomized to the control group received one text message with the Nova Scotia Mental Health and Addictions Program website link embedded in the message. This resource was freely accessible, and this service was provided as an add-on to participants' usual care. The website offers free, evidence-based e-mental health resources designed to address various psychiatric conditions with programs such as the early psychosis program, the eating disorder program, and recovery support [21]. Participants randomized to the Text4Support group received free daily, unilateral, CBT-based text messages tailored to their specific diagnosis as an add-on program to their usual care for 6 months. The supportive text messages were developed using evidence-based principles of CBT. Each message, limited to 160 characters, was designed to provide brief, accessible prompts that encourage adaptive coping, cognitive reframing, and emotional regulation. The content was collaboratively created by CBT therapists, mental health professionals, and individuals with lived experience to ensure both clinical relevance and user-centered design. The underlying mechanism is grounded in the concept that regular exposure to concise, positively framed messages can reinforce helpful thinking patterns, enhance self-efficacy, and foster emotional resilience over time. Through consistent daily reinforcement, these messages aim to augment usual care by promoting psychological well-being and perceived connectedness. Participants completed the baseline surveys on enrollment via the online link. A text message with follow-up surveys, which includes the satisfaction survey, was sent to all participants at 6 weeks, 3 months, and 6 months.

Outcome Measures

The primary outcome measure was the participants' overall satisfaction with the Text4Support program or the Nova Scotia Health e-mental health online resources. Overall satisfaction was rated on an 11-point Likert scale, where zero indicated “very dissatisfied,” 5 indicated “neutral,” and 10 showed “very satisfied.”

Secondary outcomes included the perceived impact and feedback on the Text4Support program and the Nova Scotia Health e-mental health online resources among the Text4Support and control groups. Study participants provided feedback on their satisfaction by responding to questions that assessed their engagement and use of each intervention and its perceived impact.

Engagement and use of each intervention were measured on a 5-point Likert scale (always, mostly, sometimes, rarely, and never). The perceived impact of each intervention on how participants coped with stress and loneliness, improved quality of life, overall physical and mental well-being, and the intervention's relevance, encouragement, and supportiveness were measured on a 5-point Likert scale, ranging from “strongly agree” to “strongly disagree” and including “neutral.” For analysis purposes, responses were collapsed into 3 categories: strongly agree/agree, neutral, and strongly disagree/disagree.

The reliability and validity of this scale for testing satisfaction have not been examined; however, the survey was adopted in various studies to assess user satisfaction [8,10,22].

Data Analyses

Data were analyzed using SPSS for Windows (version 28; IBM Corporation) [23]. Participants' satisfaction data were presented as a continuous variable. We measured participants' overall satisfaction by rating the program from 0 to 10 (0=very dissatisfied, 5=neutral, and 10=very satisfied). The results were presented as means and SDs. Participants' engagement and use of each intervention (Text4Support/online resources) and the perceived impact of each intervention were summarized and presented as categorical variables and reported as frequency and percentages among the control group and Text4Support group. A chi-square and Fisher exact test were run to examine any differences in reporting user satisfaction, engagement, or perception of the Text4Support or online resources among the 2 groups, and a 2-tailed criterion ($\alpha<.05$) was used to determine statistical differences. To handle missing data, we adopted imputation techniques, focusing on the last observation carried forward method.

Results

Figure 1 illustrates the flowchart of the participants. Eight hundred and ninety-eight patients were assessed for eligibility to enter the trial, of whom 781 eligible patients were randomized: 69 in the control group and 130 in the Text4Support group completed the satisfaction survey. The response rate in the control group was 18% (69/387), whereas the Text4Support group achieved a higher response rate of 33% (130/394).

Figure 1. Participant flowchart.

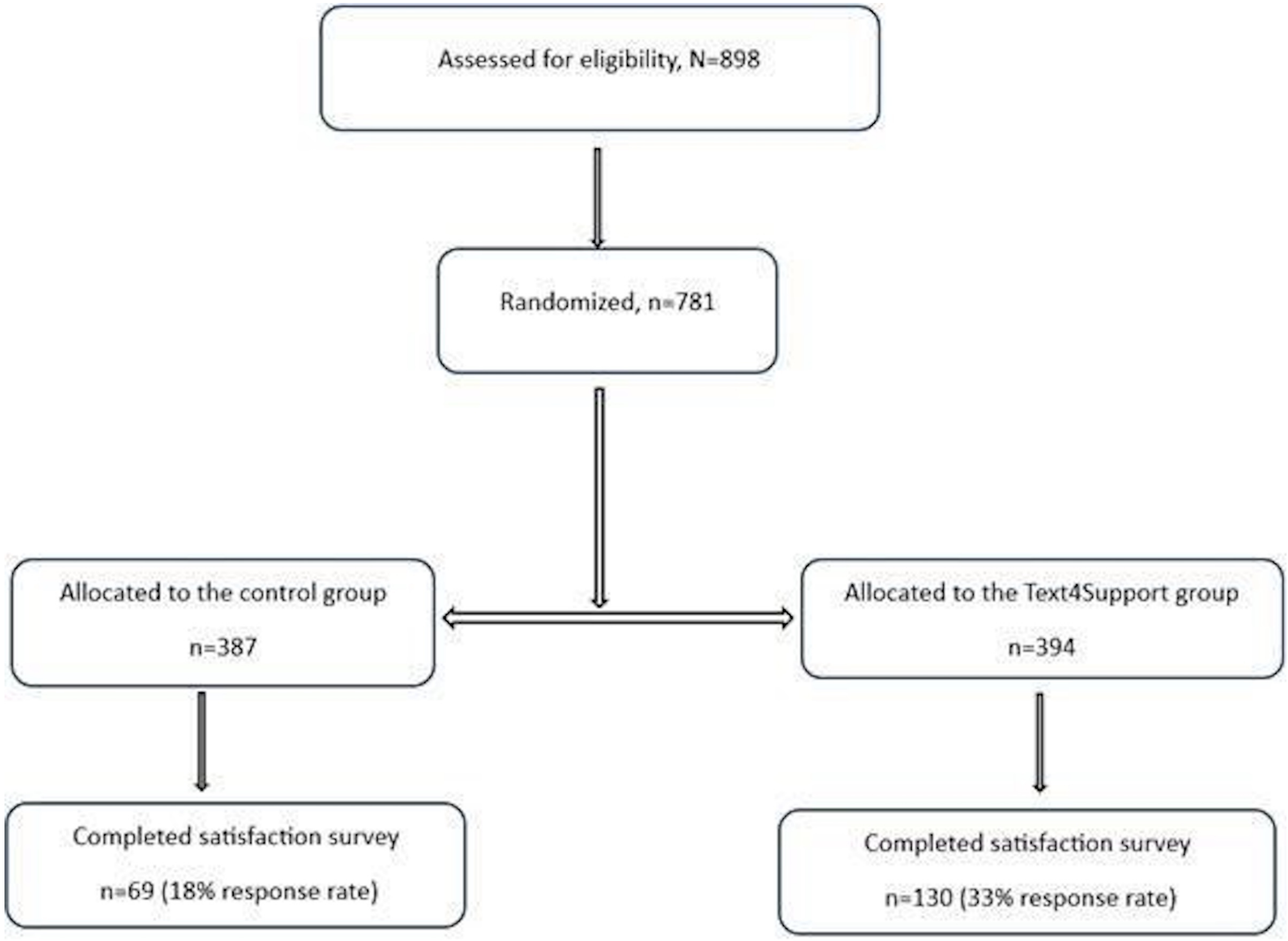


Table 1 shows the demographic characteristics of participants who completed the satisfaction survey. In total, across the 2 groups, the majority were between 26 and 60 years old (72.9%, 145/199), female (66.8%, 133/199), Caucasian (84.8%, 168/198), employed (52.3%, 103/197), and partnered/married (45.2%, 90/199) and had obtained a postsecondary education at a college or university (57.1%, 113/198). About 49% (97/199) of the study participants live in rented accommodation, 28.5% (55/193) have an income range of less than $29,590 (1 CAD≈US $0.77 at the time of the study), and 23.1% (27/117) have a primary diagnosis of depression. The chi-square analysis revealed no statistically significant differences in the demographic characteristics of participants in the Text4Support group and the control group, suggesting no differences in their demographic characteristics.

Table . Demographic characteristics of participants who completed the satisfaction survey.

Variable	Control group, n (%)	Text4Support group, n (%)	Total, n (%)	χ^2/Fisher exact *(df)*	*P* value
Age (y)					
≤25	13 (18.8)	21 (16.2)	34 (17.1)	3.1 (3)	.40
26 - 40	29 (42.0)	44 (33.8)	73 (36.7)		
41 - 60	23 (33.3)	49 (37.7)	72 (36.2)		
>60	4 (5.8)	16 (12.3)	20 (10.1)		
Sex at birth					
Male	21 (30.4)	45 (34.6)	66 (33.2)	0.4 (1)	.60
Female	48 (69.6)	85 (65.4)	133 (66.8)		
Ethnicity					
Caucasian	54 (78.3)	114 (88.4)	168 (84.8)	6.0[a]	.10
Indigenous	8 (11.6)	4 (3.1)	12 (6.1)		
Black/Hispanic	4 (5.8)	7 (5.4)	11 (5.6)		
Asian	3 (4.3)	4 (3.1)	7 (3.5)		
Employment status					
Student	7 (10.1)	11 (8.6)	18 (9.1)	2.3 (4)	.70
Employed	35 (50.7)	68 (53.1)	103 (52.3)		
Unemployed	16 (23.2)	22 (17.2)	38 (19.3)		
Retired	7 (10.1)	13 (10.2)	20 (10.2)		
Other	4 (5.8)	14 (10.9)	18 (9.1)		
Educational status					
Elementary school	2 (2.9)	2 (1.6)	4 (2.0)	1.9[a]	.80
High school	20 (29.0)	39 (30.2)	59 (29.8)		
Postsecondary (college/university)	41 (59.4)	72 (55.8)	113 (57.1)		
Postsecondary (trade school)	6 (8.7)	13(10.1)	19 (9.6)		
Other	0 (0.0)	3 (2.3)	3 (1.5)		
Relationship status					
Single	31 (44.9)	54 (42.5)	85 (42.7)	2.7[a]	.60
Partnered/married	32 (46.4)	58 (44.6)	90 (45.2)		
Separated/divorced	4 (5.8)	13 (10.0)	17 (8.5)		
Widowed	2 (2.9)	2 (1.5)	4 (2.0)		
Other	0 (0.0)	3 (2.3)	3 (1.5)		
Housing status					
Own home	19 (27.5)	43 (33.1)	62 (31.2)	1.7 (2)	.40
Rented accommodation	38 (55.1)	59 (45.4)	97 (48.7)		
Live with family/friends	12 (17.4)	28 (21.5)	40 (20.1)		
Income range (stats Canada) (1 CAD≈US $0.77 at the time of the study)					

Variable	Control group, n (%)	Text4Support group, n (%)	Total, n (%)	χ^2/Fisher exact *(df)*	*P* value
Less than $29,590	21 (32.3)	34 (26.6)	55 (28.5)	6.1 (5)	.30
$29,591-$59,180	22 (33.8)	28 (21.9)	50 (25.9)		
$59,181-$93,000	10 (15.4)	35 (27.3)	45 (23.3)		
$93,001-$150,000	5 (7.7)	12 (9.4)	17 (8.8)		
More than $150,000	4 (6.2)	10 (7.8)	14 (7.3)		
No income	3 (4.6)	9 (7.0)	12 (6.2)		
Primary diagnosis					
Depression	7 (15.2)	20 (28.2)	27 (23.1)	8.0[a]	.30
Anxiety	6 (13.0)	17 (23.9)	23 (19.7)		
Bipolar disorder	9 (19.6)	7 (9.9)	16 (13.7)		
Substance use disorder	2 (4.3)	3 (4.2)	5 (4.3)		
Alcohol use disorder	2 (4.3)	3 (4.2)	5 (4.3)		
Trauma-related disorders	9 (19.6)	7 (9.9)	16 (13.7)		
Psychosis	3 (6.5)	4 (5.6)	7 (6.0)		
Other	8 (17.4)	10 (14.1)	18 (15.4)		

[a] Fisher exact values.

The overall mean (SD) satisfaction score in the control group was 5.1 (2.3), and the overall mean (SD) satisfaction score for the Text4Support group was 7.1 (2.2).

Table 2 outlines participants' engagement with the Text4Support program and online mental health resources and the extent to which participants strongly agreed, agreed, were neutral, strongly disagreed, or disagreed with statements regarding these 2 resources. The majority of participants in the control group rarely read the eHealth resources (27/69, 39.1%), and 53.8% (70/130) in the Text4Support group always read the messages; 33.3% (23/69) in the control group sometimes takes a positive action after reading the online resources, and it was 42.3% (55/130) in the Text4Support group.

Table . Participants' feedback and perception of the supportive text messages and eHealth resources.

Perception and feedback	Control group, n (%)	Text4Support group, n (%)	Total, n (%)	χ^2/Fisher exact *(df)*	*P* value
Reading the messages/eHealth resources					
Always	6 (8.7)	70 (53.8)	76 (38.2)	79.8 (4)	<.001
Mostly	5 (7.2)	33 (25.4)	38 (19.1)		
Sometimes	19 (27.5)	16 (12.3)	35 (17.6)		
Rarely	27 (39.1)	6 (4.6)	33 (16.6)		
Never	12 (17.4)	5 (3.8)	17 (8.5)		
Read the messages/eHealth resources and take positive action after					
Always	3 (4.3)	13 (10.0)	16 (8.0)	15.5[a]	.003
Mostly	7 (10.1)	26 (20.0)	33 (16.6)		
Sometimes	23 (33.3)	55 (42.3)	78 (39.2)		
Rarely	21 (30.4)	28 21.5)	49 (24.6)		
Never	15 (21.7)	8 (6.2)	23 (11.6)		
The messages/eHealth resources were on the topic					
Strongly agree/agree	30 (46.2)	105 (81.4)	135 (69.6)	26.4[a]	<.001
Neutral	32 (49.2)	19 (14.7)	51 (26.3)		
Strongly disagree/disagree	3 (4.6)	5 (3.9)	8 (4.1)		
The messages/eHealth resources were on point					
Strongly agree/agree	29 (44.6)	104 (80.6)	133 (68.6)	25.7[a]	<.001
Neutral	32 (49.2)	21 (16.3)	53 (27.3)		
Strongly disagree/disagree	4 (6.2)	4 (3.1)	8 (4.1)		
The messages/eHealth resources were supportive					
Strongly agree/agree	27 (41.5)	105 (81.4)	132 (68.0)	31.7 (2)	<.001
Neutral	30 (46.2)	18 (14.0)	48 (24.7)		
Strongly disagree/disagree	8 (12.3)	6 (4.7)	14 (7.2)		
The messages/eHealth resources were positive					
Strongly agree/agree	32 (49.2)	114 (88.4)	146 (75.3)	38.0[a]	<.001
Neutral	30 (46.2)	10 (7.8)	40 (20.6)		
Strongly disagree/disagree	3 (4.6)	5 (3.9)	8 (4.1)		
The messages/eHealth resources improved my quality of life					

Perception and feedback	Control group, n (%)	Text4Support group, n (%)	Total, n (%)	χ^2/Fisher exact (*df*)	*P* value
Strongly agree/agree	8 (11.9)	47 (36.4)	55 (28.1)	13.7 (2)	.001
Neutral	43 (64.2)	55 (42.6)	98 (50.0)		
Strongly disagree/disagree	16 (23.9)	27 (20.9)	43 (21.9)		
The messages/eHealth resources improved my overall mental well-being					
Strongly agree/agree	12 (17.9)	66 (51.2)	78 (39.8)	22.6 (2)	<.001
Neutral	40 (59.7)	38 (29.5)	78 (39.8)		
Strongly disagree/disagree	15 (22.4)	25 (19.4)	40 (20.4)		
The messages/eHealth resources improved my overall physical well-being					
Strongly agree/agree	7 (10.4)	28 (21.7)	35 (17.9)	4.5 (2)	>.99
Neutral	40 (59.7)	61 (47.3)	101 (51.5)		
Strongly disagree/disagree	20 (29.9)	40 (31.0)	60 (30.6)		
The messages/eHealth resources helped me cope with loneliness					
Strongly agree/agree	9 (13.4)	52 (40.3)	61 (31.1)	16.3 (2)	<.001
Neutral	38 (56.7)	43 (33.3)	81 (41.3)		
Strongly disagree/disagree	20 (29.9)	34 (26.4)	54 (27.6)		
The messages/eHealth resources helped me cope with stress					
Strongly agree/agree	8 (11.9)	57 (44.2)	65 (33.2)	20.8 (2)	<.001
Neutral	40 (59.7)	47 (36.4)	87 (44.4)		
Strongly disagree/disagree	19 (28.4)	25 (19.4)	44 (22.4)		
The messages/eHealth resources were not relevant to my concern					
Strongly agree/agree	13 (20.0)	12 (9.3)	25 (12.9)	24.4 (2)	<.001
Neutral	33 (50.8)	31 (24.0)	64 (33.0)		
Strongly disagree/disagree	19 (29.2)	86 (66.7)	105 (54.1)		
The messages/eHealth resources were not helpful					
Strongly agree/agree	8 (12.3)	9 (7.0)	17 (8.8)	31.4 (2)	<.001
Neutral	34 (52.3)	22 (17.1)	56 (28.6)		
Strongly disagree/disagree	23 (35.4)	98 (76.0)	121 (62.4)		

Perception and feedback	Control group, n (%)	Text4Support group, n (%)	Total, n (%)	χ^2/Fisher exact (*df*)	*P* value
The messages/eHealth resources were not encouraging					
Strongly agree/agree	11 (16.9)	11 (8.5)	22 (11.3)	40.9 (2)	<.001
Neutral	33 (50.8)	17 (13.2)	50 (25.8)		
Strongly disagree/disagree	21 (32.3)	101 (78.3)	122 (62.9)		
The messages/eHealth resources were negative					
Strongly agree/agree	4 (6.2)	2 (1.6)	6 (3.1)	47.4[a]	<.001
Neutral	36 (55.4)	15 (11.6)	51 (26.3)		
Strongly disagree/disagree	25 (38.5)	112 (86.8)	137 (70.6)		

[a] Fisher exact values.

Most study participants were neutral when asked if the eHealth resources were on the topic (32/65, 49.2%), on point (32/65, 49.2%), supportive (30/65, 46.2%), helpful in coping with loneliness (38/67, 56.7%), and helpful in coping with stress or improving their physical and mental well-being (40/67, 59.7%). However, 49.2% (32/65) strongly agreed or agreed that the eHealth resources were positive. In terms of how eHealth improved quality of life, 64.2% (43/67) were neutral as well as for negativity of the online resources (36/67, 55.4%), not encouraging (33/65, 50.8%), not helpful (34/65, 52.3%), and not relevant (33/67, 50.8%). In contrast, the majority of participants in the Text4Support group strongly agreed or agreed that the messages were on the topic (105/129, 81.4%), were on point (104/129, 80.6%), supportive (105/129, 81.4%), positive (114/129, 88.8%), and helped them cope with loneliness (52/129, 40.3%), stress (57/129, 44.2%), or improve their overall mental well-being (66/129, 51.2%). In terms of how the messages improved quality of life, 42.6% (55/129) in the Text4Support group were neutral and strongly disagreed or disagreed with the messages being negative (112/129, 86.8%), not encouraging (101/129, 78.3%), not helpful (98/129, 76%), and not relevant (86/129, 66.7%). The results revealed statistically significant differences between groups, indicating that the observed effects were not due to chance and were not coincidental.

Discussion

Principal Findings

This study provides new insights into how patients perceive text message interventions, eHealth resources, and their engagement with the Text4Support program and the Nova Scotia Health online resources. Findings reveal that the overall mean (SD) satisfaction score in the control group was 5.1 (2.3), and the overall mean (SD) satisfaction score for the Text4Support group was 7.1 (2.2). Compared to the control group, participants in the Text4Support group reported greater engagement and positive program impact. While 53.8% (70/130) of Text4Support recipients always read the messages, only 8.7% (6/69) of the control group always read the eHealth resources. Participants allocated to the Text4Support group were reported to sometimes take action upon reading the messages (42.3% vs 33.3%). A significantly higher proportion of Text4Support users strongly agreed or agreed that the messages were supportive (81.4% vs 41.5%), positive (88.4% vs 49.2%), and helpful in coping with stress (44.2% vs 11.9%), loneliness (40.3% vs 13.4%), and improving mental well-being (51.2% vs 17.9%). In contrast, responses from the control group were largely neutral. The majority of respondents in the Text4Support group agreed or strongly agreed that the messages had a positive impact on them, which achieved statistical significance (*P*<.001) when compared with the e-mental health resource group; however, physical well-being was neutral across the 2 groups and not statistically significant (*P*=.1).

The response rates observed in this study provide important insights into participant engagement with digital mental health interventions. The Text4Support group demonstrated a notably higher response rate (33%) compared to the control group (18%). Although the response rates remain relatively low, raising concerns about potential attrition bias and the representativeness of the findings, the differences suggest that supportive SMS text messaging interventions may play a significant role in enhancing participant engagement and willingness to interact with mental health research and services. Improved response rates have important implications for the validity and generalizability of study findings. Higher response rates enhance the appropriateness and representativeness of the sample [24].

The findings indicate a high level of satisfaction among participants in the Text4Support group, with an average score of 7.1. Although our study participants expressed high satisfaction, other studies recorded higher satisfaction with the supportive text intervention [22,25-27]. For example, a study conducted among educators showed a mean satisfaction of about 8.5 among administrative educators [26]. Again, a cross-sectional study among 2032 participants recorded 8.55

overall mean satisfaction with the Text4Hope program [27]. Among the control group, an overall average satisfaction rate of 5.1 for online mental health resources suggests moderate user satisfaction. In terms of context and use, this measure aligns with the literature in that satisfaction with online mental health services is low to moderate [28]. As an example, one study making use of the Client Satisfaction Questionnaire-8 found that most patients reported low-to-moderate satisfaction with online psychiatric services [28].

The Text4Support program produced significant participant engagement, as 53% of participants consistently read text messages, while only 8.7% of control group members always read online resources. Outcomes from various studies have shown that interactive text-based interventions help maintain long-term therapeutic engagement, which delivers essential immediate mental health support [22,27,29,30]. A systematic review and meta-analysis found that SMS text messaging interventions for health promotion had widespread acceptance among participants and demonstrated high engagement rates throughout different health areas [29,30]. Although few participants always read the messages/eHealth resources and took action afterward, 73.8% and 83.8% of participants in the control and intervention groups, respectively, at least rarely read and took action in our study. However, the relatively low percentage of participants taking consistent action based on the messages suggests that passive consumption may not suffice for significant behavioral change. Integrating interactive components, such as bidirectional messaging, could enhance effectiveness. A scoping review of text message interventions in adolescent mental health services found that 65% of studies involved bidirectional messaging, indicating its potential to boost engagement and outcomes [31].

On the other hand, engagement with the provided eHealth resources was lower, with only a subset of participants actively utilizing them. Research indicates that eHealth interventions often struggle with uptake unless seamlessly integrated into daily routines [32]. Barriers to eHealth engagement usually stem from issues such as lack of time, perceived irrelevance, and digital literacy challenges [33]. Unlike text messages, which require minimal effort to consume, accessing eHealth resources demands additional steps that may deter some users. The brevity of text messages (limited to 160 characters) makes them easier to process, especially for individuals experiencing mental health challenges [34]. In contrast, studies on guided digital interventions suggest that providing step-by-step instructions or integrating behavioral prompts within text messages can enhance the utilization of eHealth resources [35]. One notable finding is that while 66.7% of the Text4Support group and 29.2% in the control group found the messages relevant, 9.3% and 20%, respectively, did not. The Text4Support group exhibited a notably higher percentage of perceived relevance, indicating that Text4Support was more effectively aligned with participants' mental health needs than the control condition. A significant percentage of participants, 9.3% in the Text4Support group and 20% in the control group, indicated that the messages/online resources were not pertinent to their needs. This disparity highlights the necessity of customizing digital interventions to individual users' specific needs, circumstances, and preferences. The control group's elevated percentage of irrelevance likely indicates the generic nature of nontargeted messages, while the Text4Support group's reduced percentage implies a degree of alignment with participants' needs, although there is still potential for enhancement.

Regarding perceived relevance and utility, the results show that most Text4Support participants (81.4%) found the messages relevant to the topic. In comparison, 46.2% strongly agreed or agreed with the online resource, underscoring the importance of tailored content. Personalization has been identified as a key factor in the success of text-based interventions. Literature has highlighted that those interventions allowing rapid, personalized exchanges, akin to natural conversation, were more effective in addressing individual needs [22,36].

The purpose of Text4Support was to provide emotional support through daily messages. The Text4Support group overwhelmingly found the messages supportive (81.4%) and positive (88.4%). These findings demonstrate that digital interventions with positive reinforcement significantly improve users' mental well-being [10,22,37,38]. Research indicates that such interventions can effectively develop therapeutic connections, meeting unmet needs for mental health care [10,27]. However, the debate persists on whether supportive messaging alone yields lasting psychological improvements. Studies suggest combining SMS text messaging with other treatment modalities, such as CBTs, may enhance effectiveness [39]. The supportive text messages delivered in this study were an addition to the patient's usual care, making it more effective.

The supportive text messages positively impacted mental well-being, with 51.2% of Text4Support participants reporting improvements, compared to 17.9% in the control group. Similarly, 40.3% of the intervention group reported reduced loneliness and 44.2% experienced lower stress levels. Other studies align with our findings; for example, a similar program, Wellness4Teacher, reported that 95.4% of its respondents reported that the messages were at least often positively impactful, while 93.3% reported that the messages were affirming [26]. These findings highlight the potential of digital interventions as supplementary tools for mental health support [40,41].

Our results show substantial differences in neutral responses across all groups in terms of improving the quality of life. The highest proportion of neutral participant responses was 64.2% in the control group and 42.6% in the Text4Support group, totaling 50% among the study sample. Participants who gave neutral responses were all neither for nor against the intervention having an effect on their quality of life. The slightly larger percentage of indifferent responses from the control group indicates either a lack of participation or ineffectiveness of regular care in significantly changing the participants' assessments of life quality. This is perhaps due to a more focused and structured approach, which is likely to impress participants more than those of general care. The findings align with previous research showing that participants' opinions about digital treatments vary [42,43]. For example, Proudfoot et al [43] found that while 68% of respondents indicated considerable improvements to their mental health following digital treatments, about 25% reported neutral or unsure answers, underscoring that outcomes are determined contextually. A qualitative research study also documents instances of most respondents feeling neutral or holding dual views about digital mental health interventions [44].

Similarly, a meta-analysis of RCTs highlighted that, as a result of the differences in people's participation and expectations, neutral answers often emerge in studies involving smartphone-based mental health therapies [42].

Across groups, participants demonstrated similar disagreement percentages, with the coping of stress registering 28.4% in the eHealth resource group and 19.4% in the Text4Support group, resulting in an overall rate of 22.4%. Although digital interventions may have received more positive feedback, they were not effective for everyone, and some participants found them ineffective in stress coping [45].

However, both interventions had a weaker effect on physical well-being, with no significant differences between groups [46,47]. This finding is consistent with prior research, which has shown that text-based interventions can effectively improve mental health outcomes but have a limited influence on physical health behaviors [47,48]. Digital interventions targeting physical health require more structured behavior-change strategies, such as goal setting and reinforcement mechanisms, rather than passive messaging alone [49].

Implications for Policy, Practice, and Future Research

The findings of this study have several important implications for the design, implementation, and scaling of digital mental health interventions. First, the significantly higher satisfaction and engagement levels observed in the Text4Support group underscore the value of incorporating supportive, brief, and personalized messaging into routine mental health care. Policymakers and health care providers should consider integrating text-based interventions as a standard adjunct to traditional care, particularly for populations that may face barriers to in-person or web-based services.

Given the low engagement with static eHealth resources, health systems should prioritize low-barrier, mobile-first tools that align with users' daily routines. Moreover, to increase the effectiveness of such interventions, future policy should encourage the development of interactive, bidirectional messaging systems that foster active participation and allow for real-time support and personalization based on user feedback and needs. Tailoring content to individual needs, such as demographic or diagnostic profiles, can further enhance relevance and efficacy, as evidenced by the higher perceived relevance and positive impact of the Text4Support messages.

Although Text4Support currently relies on preprogrammed, therapist-developed messages, future iterations of supportive SMS text messaging interventions could leverage artificial intelligence, including generative artificial intelligence, to enhance message compilation, personalization, and responsiveness. Finally, given the limited impact on physical well-being, digital programs should integrate behavior change techniques (eg, goal setting, reminders, and progress tracking) when addressing physical health outcomes. Health authorities should support cross-sector collaboration between digital health developers, clinicians, and researchers to ensure interventions are both evidence-based and contextually responsive. These findings advocate for a patient-centered, adaptive approach to mental health support, where digital tools are not merely supplemental but are integral to comprehensive, accessible, and scalable care strategies.

Limitations

Our study has limitations that need to be considered when appraising the results. First, the generalizability of the results may be affected by the comparatively small sample size of participants who completed the follow-up satisfaction survey, particularly in the control group. The imbalance in group sizes (130 in the Text4Support group vs 69 in the control group) may have introduced bias or limited the statistical power to detect differences in some outcomes. The small sample size may be attributed to the online nature of the surveys; literature has reported that online distributed surveys are unlikely to keep participants in a study for follow-up assessment compared to paper-based surveys [50]. One notable limitation of this study is the relatively low survey completion rate. While this may impact the generalizability of the findings, it is consistent with patterns observed in other digital mental health and eHealth trials, which frequently report attrition rates exceeding 30% [51,52]. Again, participants' satisfaction, engagement, and perception were self-reported, which may result in recall bias since participants may overestimate their feedback. Lastly, the web-based questionnaire used to assess participant satisfaction and perception was not a validated instrument, which may reduce the reliability of the reported satisfaction rates. Nonetheless, this was one of the first RCTs examining user satisfaction among Text4Support users and Nova Scotia mental health and addiction online resources users in Nova Scotia.

Conclusions

This study's findings demonstrate the meaningful impact of personalized, supportive SMS text messaging in enhancing engagement, emotional well-being, and satisfaction among mental health service users. As mental health systems evolve to meet rising demands, digital interventions like Text4Support represent a promising, scalable solution that can complement existing care. To maximize their potential, policies must prioritize accessibility, personalization, and integration with broader care models, ensuring that digital tools meet people where they are, with the right message at the right time.

Funding

The study was funded by the Queen Elizabeth II Foundation and supported by the Global Psychological eHealth Foundation, the Department of Psychiatry at Nova Scotia Health, and Dalhousie University. The funders had no role in the study's design, execution,

data collection, analysis, interpretation, or writing. They were also not involved in reviewing, approving, or deciding to publish the results.

Data Availability

The data supporting this manuscript's conclusions will be made available upon reasonable request from the corresponding author.

Authors' Contributions

VIOA conceived, designed, and supervised the study. VIOA, RS, and GO-D conducted data analysis. The initial manuscript was drafted by GO-D. VIOA, RS, BA, SON, MKA, EE, and LW reviewed the initial draft. All the authors contributed to the study's design, revised the initial draft manuscript, and approved the final draft before submission. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

VIOA is the founder and Principal Investigator of the ResilienceNHope messaging programs and volunteers as the Board Chair of the Global Psychological eHealth Foundation, a not-for-profit organization. BA volunteers as the President and Chief Executive Officer of the Global Psychological eHealth Foundation. RS volunteers as the Program Director of the Global Psychological eHealth Foundation. VIOA, BA, and RS receive no financial or other compensation from the Global Psychological eHealth Foundation and have no financial conflicts of interest about this article. GO-D, SON, MA, EE, and LW have no commercial or financial conflict of interest.

Checklist 1
CONSORT-eHEALTH (v 1.6.1) checklist.
[PDF File, 1144 KB - humanfactors_v13i1e76587_app1.pdf]

References

1. Børge M, Yngvild A. User satisfaction with child and adolescent mental health services: the association between user satisfaction and clinical outcomes. BMC Psychiatry 2024 Apr 15;24(1):279. [doi: 10.1186/s12888-024-05715-1] [Medline: 38622603]
2. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ Open 2013 Jan 3;3(1):e001570. [doi: 10.1136/bmjopen-2012-001570] [Medline: 23293244]
3. Anhang Price R, Elliott MN, Zaslavsky AM, et al. Examining the role of patient experience surveys in measuring health care quality. Med Care Res Rev 2014 Oct;71(5):522-554. [doi: 10.1177/1077558714541480] [Medline: 25027409]
4. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013;35(4):332-338. [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
5. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
6. Aguilera A, Muñoz RF. Text messaging as an adjunct to CBT in low-income populations: a usability and feasibility pilot study. Prof Psychol Res Pr 2011 Dec 1;42(6):472-478. [doi: 10.1037/a0025499] [Medline: 25525292]
7. Obuobi-Donkor G, Shalaby R, Eboreime E, et al. Text4PTSI: a promising supportive text messaging program to mitigate psychological symptoms in public safety personnel. Int J Environ Res Public Health 2023 Feb 27;20(5):4215. [doi: 10.3390/ijerph20054215] [Medline: 36901235]
8. Agyapong VIO, Mrklas K, Juhás M, et al. Cross-sectional survey evaluating Text4Mood: mobile health program to reduce psychological treatment gap in mental healthcare in Alberta through daily supportive text messages. BMC Psychiatry 2016 Nov 8;16(1):378. [doi: 10.1186/s12888-016-1104-2] [Medline: 27821096]
9. Agyapong VIO, Hrabok M, Vuong W, et al. Changes in stress, anxiety, and depression levels of subscribers to a daily supportive text message program (Text4Hope) during the COVID-19 pandemic: cross-sectional survey study. JMIR Ment Health 2020 Dec 18;7(12):e22423. [doi: 10.2196/22423] [Medline: 33296330]
10. Agyapong B, Shalaby R, Eboreime E, et al. Subscribers' perspectives and satisfaction with the MoreGoodDays supportive text messaging program and the impact of the program on self-rated clinical measures. J Clin Med 2024 Jan 19;13(2):580. [doi: 10.3390/jcm13020580] [Medline: 38276086]
11. Kornfield R, Stamatis CA, Bhattacharjee A, 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. [doi: 10.1016/j.invent.2023.100667] [Medline: 37746639]
12. Agyapong VIO, Juhás M, Mrklas K, et al. Randomized controlled pilot trial of supportive text messaging for alcohol use disorder patients. J Subst Abuse Treat 2018 Nov;94:74-80. [doi: 10.1016/j.jsat.2018.08.014] [Medline: 30243421]

13. Agyapong VIO, Juhás M, Ohinmaa A, et al. Randomized controlled pilot trial of supportive text messages for patients with depression. BMC Psychiatry 2017 Aug 2;17(1):286. [doi: 10.1186/s12888-017-1448-2] [Medline: 28768493]
14. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for future mobile-Health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. J Med Internet Res 2016 Jun 10;18(6):e135. [doi: 10.2196/jmir.5066] [Medline: 27287668]
15. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11. [doi: 10.1002/wps.20083] [Medline: 24497236]
16. Karasouli E, Adams A. Assessing the evidence for e-Resources for mental health self-management: a systematic literature review. JMIR Ment Health 2014;1(1):e3. [doi: 10.2196/mental.3708] [Medline: 26543903]
17. 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]
18. Good clinical practices. Government of Canada. 2022. URL: https://www.canada.ca/en/health-canada/services/drugs-health-products/compliance-enforcement/good-clinical-practices.html [accessed 2025-12-24]
19. Dias RDL, Shalaby R, Agyapong B, et al. Augmenting mental health support for patients accessing different degrees of formal psychiatric care through a supportive text messaging program: protocol for a randomized controlled trial. Methods Protoc 2023 Feb 13;6(1):19. [doi: 10.3390/mps6010019] [Medline: 36827506]
20. 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. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
21. Mental health and addictions tools. Nova Scotia Mental Health and Addictions. 2025. URL: https://mha.nshealth.ca/en/tools [accessed 2025-01-01]
22. Obuobi-Donkor G, Eboreime E, Shalaby R, et al. User satisfaction with a daily supportive text message program (Text4PTSI) for public safety personnel: longitudinal cross-sectional study. JMIR Form Res 2023 Jun 23;7:e46431. [doi: 10.2196/46431] [Medline: 37351940]
23. IBM SPSS statistics for windows, version 28.0. IBM Corp. 2021. URL: https://www.ibm.com/products/spss-statistics [accessed 2025-01-01]
24. Holtom B, Baruch Y, Aguinis H, A Ballinger G. Survey response rates: trends and a validity assessment framework. Human Relations 2022 Aug;75(8):1560-1584. [doi: 10.1177/00187267211070769]
25. Cartujano-Barrera F, Sanderson Cox L, Arana-Chicas E, et al. Feasibility and acceptability of a culturally- and linguistically-adapted smoking cessation text messaging intervention for Latino smokers. Front Public Health 2020;8:269. [doi: 10.3389/fpubh.2020.00269] [Medline: 32714891]
26. Agyapong B, Brett-MacLean P, da Luz Dias R, Wei Y, Agyapong VIO. Satisfaction with a daily supportive text messaging program aimed at promoting psychological well-being of educators: Outcomes from implementation of the Wellness4Teachers. Digit Health 2025;11:20552076251317558. [doi: 10.1177/20552076251317558] [Medline: 40904755]
27. Shalaby R, Vuong W, Hrabok M, et al. Gender differences in satisfaction with a text messaging program (Text4Hope) and anticipated receptivity to technology-based health support during the COVID-19 pandemic: cross-sectional survey study. JMIR Mhealth Uhealth 2021 Apr 15;9(4):e24184. [doi: 10.2196/24184] [Medline: 33750738]
28. Jilani AQ, Khan A, Saloni S, et al. Level of patient satisfaction with online psychiatric outdoor services. Consort Psychiatr 2023 Sep 29;4(3):23-32. [doi: 10.17816/CP5597] [Medline: 38249530]
29. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32(1):56-69. [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
30. Head KJ, Noar SM, Iannarino NT, Grant Harrington N. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]
31. MacDougall S, Jerrott S, Clark S, Campbell LA, Murphy A, Wozney L. Text message interventions in adolescent mental health and addiction services: scoping review. JMIR Ment Health 2021 Jan 8;8(1):e16508. [doi: 10.2196/16508] [Medline: 33416504]
32. Free C, Phillips G, Galli L, 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. [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
33. Yardley L, Spring BJ, Riper H, 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]
34. Lim MSC, Wright C, Hellard ME. The medium and the message: fitting sound health promotion methodology into 160 characters. JMIR Mhealth Uhealth 2014 Nov 3;2(4):e40. [doi: 10.2196/mhealth.3888] [Medline: 25367387]
35. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
36. Hoermann S, McCabe KL, Milne DN, Calvo RA. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. J Med Internet Res 2017 Jul 21;19(8):e267. [doi: 10.2196/jmir.7023] [Medline: 28784594]
37. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: integrating smartphone technology in behavioral healthcare. Professional Psychology: Research and Practice 2011;42(6):505-512. [doi: 10.1037/a0024485]

38. Shalaby R, Vuong W, Eboreime E, Surood S, Greenshaw AJ, Agyapong VIO. Patients' expectations and experiences with a mental health-focused supportive text messaging program: mixed methods evaluation. JMIR Form Res 2022 Jan 11;6(1):e33438. [doi: 10.2196/33438] [Medline: 35014972]
39. Senanayake B, Wickramasinghe SI, Chatfield MD, Hansen J, Edirippulige S, Smith AC. Effectiveness of text messaging interventions for the management of depression: A systematic review and meta-analysis. J Telemed Telecare 2019 Oct;25(9):513-523. [doi: 10.1177/1357633X19875852] [Medline: 31631764]
40. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13. [doi: 10.2196/jmir.1194] [Medline: 19403466]
41. Muñoz RF, Chavira DA, Himle JA, et al. Digital apothecaries: a vision for making health care interventions accessible worldwide. Mhealth 2018;4:18. [doi: 10.21037/mhealth.2018.05.04] [Medline: 30050914]
42. Firth J, Torous J, Nicholas J, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298. [doi: 10.1002/wps.20472] [Medline: 28941113]
43. Proudfoot J, Clarke J, Birch MR, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013 Nov 18;13:312. [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
44. Lau CKY, Saad A, Camara B, Rahman D, Bolea-Alamanac B. Acceptability of digital mental health interventions for depression and anxiety: systematic review. J Med Internet Res 2024 Oct 28;26:e52609. [doi: 10.2196/52609] [Medline: 39466300]
45. Luo C, Sanger N, Singhal N, et al. A comparison of electronically-delivered and face to face cognitive behavioural therapies in depressive disorders: A systematic review and meta-analysis. EClinicalMedicine 2020 Jul;24:100442. [doi: 10.1016/j.eclinm.2020.100442] [Medline: 32775969]
46. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336. [doi: 10.1002/wps.20673] [Medline: 31496095]
47. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567. [doi: 10.2196/14567] [Medline: 31573916]
48. Smith DM, Duque L, Huffman JC, Healy BC, Celano CM. Text message interventions for physical activity: a systematic review and meta-analysis. Am J Prev Med 2020 Jan;58(1):142-151. [doi: 10.1016/j.amepre.2019.08.014] [Medline: 31759805]
49. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4. [doi: 10.2196/jmir.1376] [Medline: 20164043]
50. Daikeler J, Bošnjak M, Lozar Manfreda K. Web versus other survey modes: an updated and extended meta-analysis comparing response rates. J Surv Stat Methodol 2020 Jun 1;8(3):513-539. [doi: 10.1093/jssam/smz008]
51. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
52. Wang J, Mahe J, Huo Y, et al. Framework development for reducing attrition in digital dietary interventions: systematic review and thematic synthesis. J Med Internet Res 2024 Aug 27;26:e58735. [doi: 10.2196/58735] [Medline: 39190910]

Abbreviations

CBT: cognitive behavioral therapy
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture

Edited by A Kushniruk; submitted 27.Apr.2025; peer-reviewed by D Gratzer, TTO Kwok; revised version received 01.Dec.2025; accepted 14.Dec.2025; published 21.Jan.2026.

Please cite as:
Obuobi-Donkor G, Shalaby R, Agyapong B, Obeng Nkrumah S, Adu MK, Eboreime E, Wozney L, Agyapong VIO
Evaluating User Engagement and Satisfaction With Digital Mental Health Interventions: Randomized Controlled Trial of a Text Messaging Program and e-Mental Health Resources
JMIR Hum Factors 2026;13:e76587
URL: https://humanfactors.jmir.org/2026/1/e76587
doi:10.2196/76587

Designing a Substance Misuse Data Dashboard for Overdose Fatality Review Teams: User-Centered Design Approach

Marie Pisani[1], MD; Madeline K Oguss[2], MS; Julia Dickson-Gomez[3], PhD; Constance Kostelac[3], PhD; Amy Parry[3], MPH; Starr Moss[4], MS; Elizabeth Salisbury-Afshar[5], MD, MPH; Brian Patterson[6], MD, MPH; Michael Spigner[6], MD, NRP; Megan Gussick[6], MD; Alison Krautkramer[7], MS; Timothy Gruenloh[7], MS; Askar Safipour Afshar[2], MS; Preeti Gupta[2,8], MD, MPH; Anoop Mayampurath[2,7], PhD; Majid Afshar[2,7], MD, MS

[1] Department of Medicine, University of Wisconsin, Madison, WI, United States
[2] School of Medicine and Public Health, University of Wisconsin, 610 Walnut St, Suite 517, Madison, WI, United States
[3] Institute for Health and Equity, Medical College of Wisconsin, Milwaukee, WI, United States
[4] United States Department of Justice, Madison, WI, United States
[5] Department of Family Medicine and Community Health, University of Wisconsin, Madison, WI, United States
[6] BerbeeWalsh Department of Emergency Medicine, University of Wisconsin, Madison, WI, United States
[7] Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, WI, United States
[8] Division of Pulmonary, Critical Care, Sleep, and Allergy, University of Illinois Chicago, Chicago, IL, United States

Corresponding Author:
Majid Afshar, MD, MS
School of Medicine and Public Health, University of Wisconsin, 610 Walnut St, Suite 517, Madison, WI, United States

Abstract

Background: Overdose fatality review (OFR) is a public health process in which cases of fatal overdose are carefully reviewed to identify prevention strategies. Current OFR requires review of multiple unconnected data sources, which is a manually intensive process. The Substance Misuse Data Commons (SMDC) was created to link electronic health record data with data from local and state agencies into a single, cloud-based e-platform but does not currently have a data visualization tool.

Objective: We aimed to use human factors design principles to develop a comprehensive dashboard for the SMDC that could facilitate enhanced processes to support OFR.

Methods: We first surveyed OFR leaders in Wisconsin using the National Aeronautics and Space Administration-Task Load Index to understand the cognitive workload of 3 tasks: (1) analysis of population-level overdose trends, (2) selection and preparation of individual cases for review, and (3) abstraction of data from individual causes. We then conducted semistructured interviews to identify targets for workflow optimization. Next, we developed a prototype dashboard for evaluation using a synthetic dataset built with GPT-4. We subsequently performed iterative design sessions with heuristic evaluations and collected end-user feedback on the final prototype via a second round of semistructured interviews and targeted surveys, including the Unified Theory of Acceptance and Use of Technology and the Perceived Usefulness Questionnaire.

Results: The National Aeronautics and Space Administration-Task Load Index revealed a moderately high mental workload with the current workflow for all 3 tasks, with mean scores of 12.60 (SD 3.31), 11.90 (SD 3.57), and 12.43 (SD 5.41) for tasks 1, 2, and 3, respectively. Interviews pointed to causes including technological challenges and a reliance on manual processes. The prototype dashboard addressed these concerns by integrating multiple data sources to generate population-level visualizations and patient-level event timelines. End users reported the potential for improved efficiency and data accessibility compared to antecedent processes. The Unified Theory of Acceptance and Use of Technology results indicated the dashboard would likely be adopted if made available, with a mean of 4.07 out of 5.00 (SD 0.65). The Perceived Usefulness Questionnaire results suggested moderate usefulness for both the aggregate and individual-level data, with means of 3.61 (SD 0.82) and 3.64 (SD 0.85) out of 5.00, respectively.

Conclusions: OFR is a data-intensive process that traditionally demands substantial cognitive and manual effort, and there are multiple barriers to efficiently collecting data and presenting them for review. The dashboard offers a user-centered, informatics-based approach to streamline data aggregation and presentation, potentially enhancing the efficiency of case reviews. Implementing a dashboard that consolidates and visualizes disparate data sources has the potential to alleviate the manual workload in OFR. Ultimately, our aim is to deliver a finalized data dashboard with real-world SMDC data, giving OFR leaders additional tools to aid in their rigorous work shaping interventions to reduce overdose fatalities.

(JMIR Hum Factors 2026;13:e79407) doi:10.2196/79407

KEYWORDS

data linkage; overdose; public health informatics; qualitative research; substance use disorders

Introduction

Drug overdose death rates have been steadily rising over the past 2 decades, with the most significant annual spike occurring between 2019 and 2020, when rates surged by 31.0% in the United States [1]. To address this public health crisis, some local health departments around the country have assembled overdose fatality review (OFR) teams. OFR teams are multidisciplinary and multiagency teams with representatives from areas such as public health, safety, social services, medical examiners and coroners, emergency responders, substance use treatment providers, and other community stakeholders [2]. These teams discuss the local trends in overdose fatalities and review individual cases to identify and implement recommendations aimed at preventing overdose deaths [2]. OFR teams can help facilitate harm reduction strategies such as syringe services programs, naloxone education and training, outreach programs, and coordination of treatment services from health care settings [3]. However, many counties have yet to implement OFRs, and those that exist exhibit considerable heterogeneity in their data collection procedures, requiring substantial effort in data curation.

OFR teams in Wisconsin perform fatality reviews by integrating population- and individual-level data. These reviews involve an in-depth exploration of an individual's timeline before their death, focusing on potential opportunities for intervention. The data may come from different sources, including the medical examiners and coroners office, emergency medical services (EMS), the Department of Corrections, law enforcement, social media, and other local and state agencies. Often, obtaining these data requires collecting data on individual cases from each agency and linking across multiple sources.

The Substance Misuse Data Commons (SMDC) is a single, cloud-based data repository that links hospital electronic health record (EHR) data for patients with substance misuse to local and state agency data from EMS, the Department of Corrections, the Prescription Drug Monitoring Program, state and national death sources, statewide medical and pharmacy claims, and neighborhood-level socioeconomic data [4]. It was recently created by our team to address the issue of siloed datasets for substance misuse research, but it currently lacks an interface for data visualization [4]. Dashboards have been recognized as effective tools for visualizing public health data and facilitating disease surveillance, targeted analyses, and decision-making [5]. Additionally, a well-designed dashboard can reduce cognitive workload and improve efficiency by decreasing the amount of time spent gathering data [6]. In several states, such as Indiana and North Carolina, health departments have used dashboards of aggregate data to monitor overdose trends and uncover preventable risk factors [7,8]. While these dashboards are effective for tracking local and state-wide trends and setting case review priorities, they often lack the breadth of data and the level of detail needed by OFR teams. Currently, OFR teams in Wisconsin do not have access to the SMDC. Their data collection processes rely on manually requesting and compiling information from multiple partner agencies. The SMDC was established as a research infrastructure to enable secure multiagency data linkage, and this study represents the first effort to design a visualization interface that could make those linked data accessible and actionable for OFR teams in the future.

The objective of Phase 1 of this study was to assess the cognitive workload of OFR teams and their currently utilized data processes. The objective of Phase 2 of this study was to design a prototype data dashboard and then assess the usability and acceptability of the dashboard in a simulated study. This study introduces a novel, user-centered approach to understanding and improving the OFR process by combining cognitive workload assessment with prototype dashboard design. Unlike prior public health surveillance dashboards, our approach integrates multiagency data within a unified framework modeled on real OFR workflows and leverages synthetic data generation using large language models to enable privacy-preserving development and testing.

Methods

Phase 1: Analysis of Current Workflow

We conducted surveys and semistructured interviews with OFR leaders in Wisconsin to document the current workflow and associated cognitive workload for data collection and presentation. Participants were eligible if they took part in at least one of 3 tasks: (1) aggregating or analyzing population-level data, (2) selecting cases for case reviews, or (3) abstracting data for individual case reviews. Of the 29 counties with OFR teams, we recruited participants from 2 counties—one corresponding to the dataset currently included in our study, and the other in anticipation of future data collection for the SMDC. Recruitment occurred from organizational listservs in response to informational emails sent by local OFR leadership, followed by an invitation to the voluntary and confidential survey. Participants provided consent and were enrolled via a link to the survey in a secure, web-based Research Electronic Data Capture (REDCap; Vanderbilt University) database [9,10].

We used the National Aeronautics and Space Administration-Task Load Index (NASA-TLX) survey to assess the cognitive workload of the contemporary OFR process for each task listed above that the OFR leader participated in [11]. The NASA-TLX is a validated, multidimensional tool that assesses a task's subjective cognitive workload across 6 dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration [11]. The survey asks participants to rate each dimension on a scale between 1 and 20, with higher scores indicating greater perceived workload. For the performance dimension, lower scores correspond to higher perceived success [11]. Example items include "How mentally demanding was the task?" and "How hurried or rushed was the pace of the task?" Participants rated each applicable task they performed during the OFR process.

Domain-specific mean scores were calculated rather than a composite overall score, as the NASA-TLX dimensions are designed to capture distinct aspects of workload rather than a single latent construct. Therefore, the mean and SD for each dimension were calculated.

Survey participants were then invited to a 30-minute semistructured interview with a research team member (MP). This interview approach combined a predefined structure with flexibility for follow-up questions based on participants' responses [12]. The interview guide focused on participants' responsibilities, workflows, mental workload, key data sources, and challenges. Survey findings were shared, and participants were asked to provide insight on the results. Wisconsin OFR training and technical assistance providers and a senior qualitative research expert (JDG) reviewed the guide to ensure its relevance and rigor. Interviews were conducted between October 2023 and January 2024 with participants providing verbal informed consent through secure, virtual conference software. The audio recording was transcribed and reviewed by the interviewer.

Interviews were stored and analyzed using MAXQDA 2024 (VERBI Software). The constant comparative method was used to analyze the transcripts, and inductive coding was applied to organize the information into emergent categories [13]. The codebook was repeatedly revisited and revised during the process [13]. When no new categories emerged from the analysis of additional interviews, code saturation was determined to be met. Using the final codebook, all interviews were coded by 1 researcher. Around 3 interviews were coded by a second researcher to assess intercoder reliability (ICR), exceeding the typical 10% - 25% double-coded interviews required to establish ICR [14]. The 2 researchers then compared coded segments, and disagreements were adjudicated.

Phase 2: Data Dashboard Design and Evaluation

To mitigate data privacy concerns during dashboard development, a synthetic dataset was patterned on SDMC data [15]. The synthetic data were generated using the GPT-4 application programming interface (OpenAI) [16], with chain-of-thought instructions to create each variable from the data dictionary, similar to other best practices in prompt engineering [17]. The prompt incorporated aggregate cohort demographic descriptive statistics to preserve variable distributions within the SMDC dataset. Initial prompts were tested with 20 rows of patient data to evaluate the quality of the output before a dataset of 300 patients was created. This dataset was then scaled up to 1273 patients using YData Fabric (YData AI), a synthetic data generation software that employs generative adversarial networks to produce large volumes of data accurately replicating the statistical characteristics of the original data—in this case, the smaller synthetic dataset [18].

An initial prototype of the dashboard was developed using Microsoft Power Business Intelligence (Microsoft 2024) [19]. The dashboard was designed to emulate the workflow and highlight the priority data sources identified in Phase 1. The dashboard consisted of 3 functional components: (1) visualizations of population-level data to identify demographics and trends (Figure 1); (2) line-level data, such as individual patient timelines, to facilitate case-based reviews (Figure 2); and (3) prediction tools, including census tract-level EMS patient incident predictions, deidentified hospital note topics, and a 30-day risk score for hospital readmission or death.

We conducted 2 iterative design sessions with emergency medicine physicians, including 2 EMS medical directors, a prehospital informatician, and a clinical human factors design expert. This team provided expertise in data visualization principles and linking hospital systems with prehospital emergency services, while allowing us to reserve OFR leaders for participation in the main study. The Heuristic Evaluation Checklist for Dashboard Visualizations was used to identify and address major usability issues [20]. These sessions refined the dashboard's content, organization, and visual elements, culminating in a final prototype. After finalizing the prototype, a demonstration video was shared with participants to showcase its content, organization, and key features.

Figure 1. Substance misuse data dashboard: population-level visualizations.

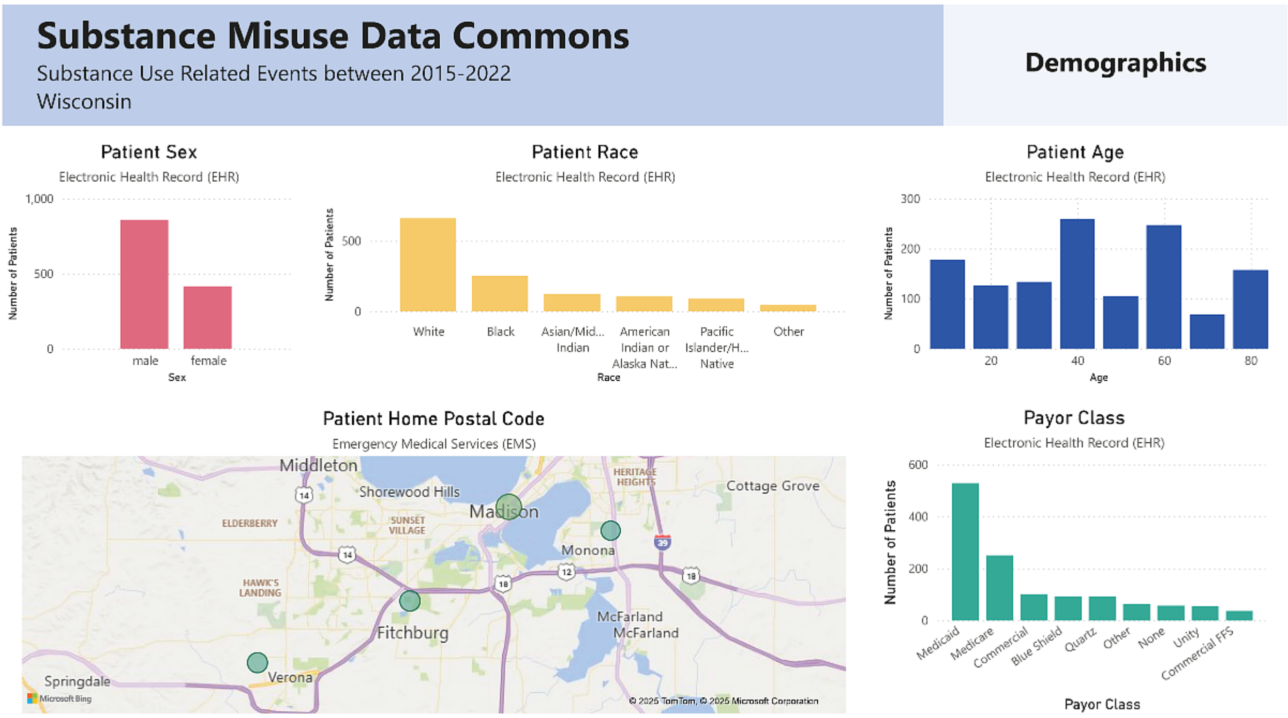
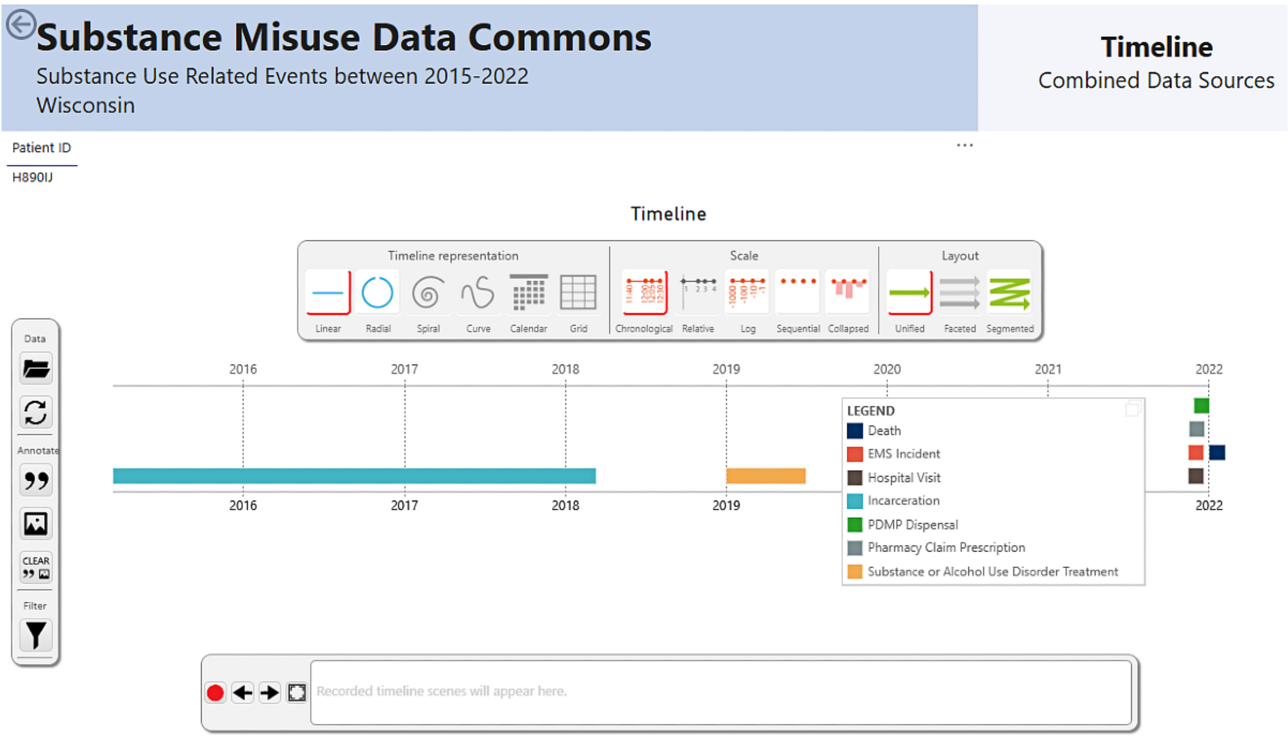


Figure 2. Substance misuse data dashboard: patient timeline visualizations.



Phase 2: Dashboard Design and Evaluation

End-user perceptions of the final prototype were assessed with semistructured interviews and surveys, using the same strategy and procedures to recruit OFR leaders as Phase 1. A total of 2 validated survey tools were used to evaluate end-user probability of adoption (Unified Theory of Acceptance and Use of Technology [UTAUT]) [21] and perception of usefulness (Perceived Usefulness Questionnaire) [22]. The interview guide explored participants' perceptions of the dashboard, conditions for use, advantages, and suggested changes. One researcher coded all 7 interviews using the final codebook, and a second researcher independently coded 3 interviews to establish ICR. Disagreements were adjudicated. The mean and SD of all survey responses were calculated.

The full set of NASA-TLX, UTAUT, and Perceived Usefulness Questionnaire questions; our interview guides and codebooks; our synthetic dataset and prompts; and our demonstration video are all viewable in our GitLab repository [23].

Ethical Considerations

The study followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guidelines [24] (Checklist 1). This research was reviewed and approved by the Institutional Review Board at the University of Wisconsin, Madison (Institutional Review Board number 2023‐1091). Informed consent verbiage was included in our recruitment emails, and verbal consent was obtained during interviews. Participant email addresses were collected in the survey in order to invite the individual to participate in an interview. Survey and interview results were deidentified. No compensation was provided to participants.

Results

Phase 1: Analysis of Current Workflow

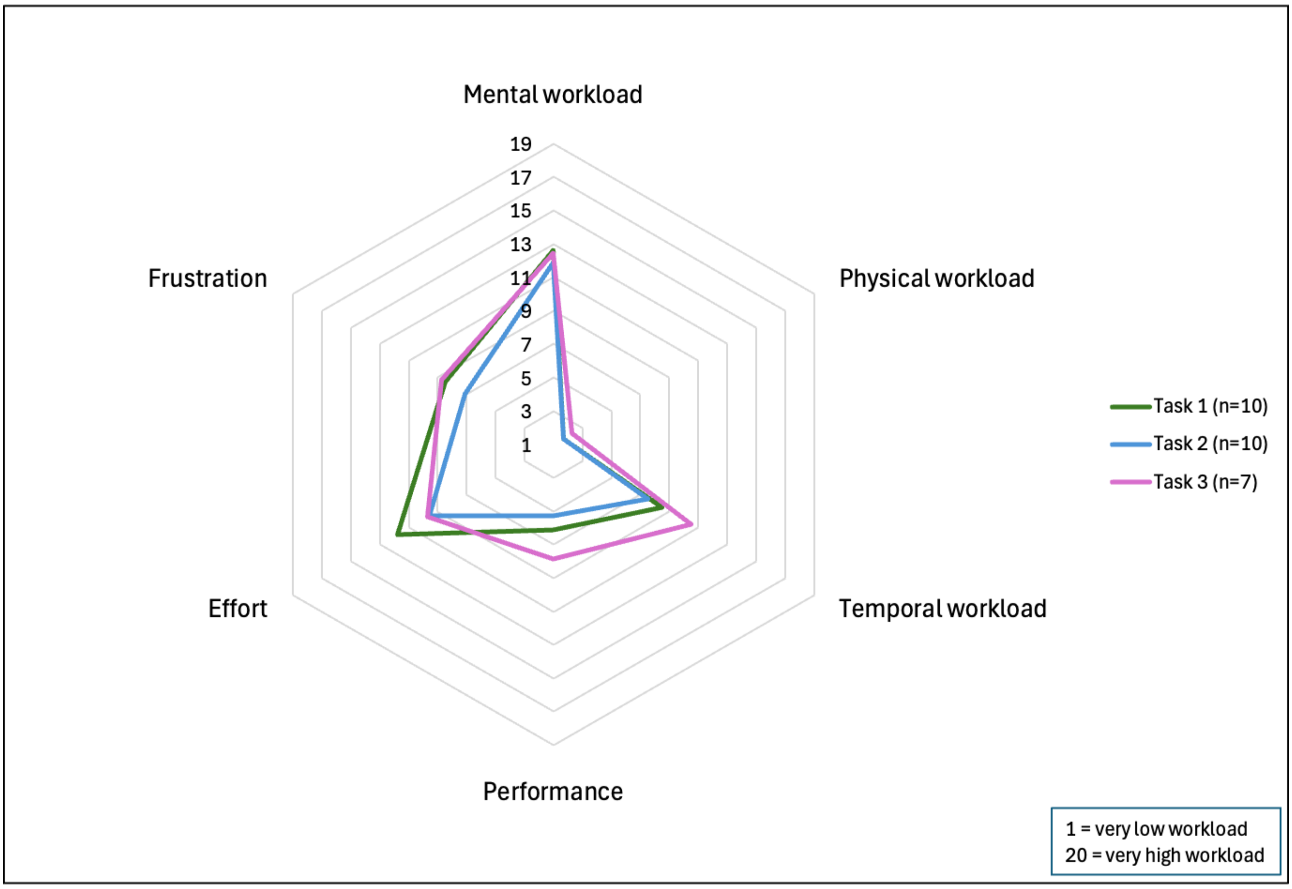
Surveys

A total of 11 OFR leaders, representing both county- and state-level agencies, participated in the survey. Table 1 provides the characteristics of the participants. The survey assessed the cognitive workload of 3 distinct tasks: (1) aggregating or analyzing population-level data, (2) selecting cases for case reviews, or (3) abstracting data for individual case reviews. Tasks had unequal sample sizes due to variations in task participation among participants. High mental workload was reported across all tasks (Figure 3). The time required for each task varied among participants. Per case review period of 1-3 months depending on the team, aggregating population-level data took an average of 5.5 hours (SD 3.09), selecting cases for case review averaged 7.95 hours (SD 5.67), and abstracting data for individual case reviews required an average of 10.50 hours (SD 8.86).

Table . Phases 1 and 2 survey demographics for overdose fatality review leaders in Wisconsin, 2023‐2024.

Demographic variables	Phase 1 participants (n=11)	Phase 2 participants (n=6)
Age (y), n (%)		
20‐29	5 (45.4)	2 (33.3)
30‐39	3 (27.3)	0 (0)
40‐49	2 (18.2)	3 (50)
50‐59	1 (9.1)	1 (16.7)
Sex, n (%)		
Female	10 (90.9)	6 (100)
Race, n (%)		
White	11 (100)	6 (100)
Education, n (%)		
Technical school, vocational training, community college	1 (9.1)	0 (0)
Bachelor’s degree	1 (9.1)	1 (16.7)
Master’s degree	9 (91.8)	5 (83.3)
Sector, n (%)		
Public health	9 (81.8)	6 (100)
Government	1 (9.1)	0 (0)
Education	1 (9.1)	0 (0)

Figure 3. National Aeronautics and Space Administration-Task Load Index (NASA-TLX) survey results for overdose fatality review leaders in Wisconsin, 2023-2024.



Interviews

Key Data Sources

In total, 10 of the 11 participants completed qualitative follow-up interviews. Participants were asked to describe their data collection process, case selection methods, challenges, desires for a future state, and key data sources. Key OFR data sources highlighted by participants are presented in Table 2. Some of these sources were used for individual or population-level data only, but many were used for both. How these sources were used varied among participants. Summaries and selected quotes are highlighted below, while the full interview results and quotes are detailed in Multimedia Appendix 1.

Table . Key data sources for overdose fatality review leaders in Wisconsin, 2023 - 2024.

Category	Data contributors and resources
High-impact data contributors	1. State vital records 2. Next-of-kin interviews 3. OFR[a] agency partners
Key agency partners and contributors	1. ME's[b] or coroner's office 2. EMS[c] 3. DOC[d] 4. Law enforcement
Other data sources	1. Prescription Drug Monitoring Program 2. Electronic Surveillance System for the Early Notification of Community-Based Epidemics 3. First Watch (fire and rescue) 4. Wisconsin Statewide Health Information Network 5. Overdose Detection Mapping Application Program 6. Consolidated Court Automation Programs 7. Internet 8. News 9. Social media

[a] OFR: overdose fatality review.

[b] ME: medical examiner.

[c] EMS: emergency medical services.

[d] DOC: Department of Corrections.

Process

Participants described a typical workflow while preparing for an OFR meeting. Population-level data are analyzed to identify recent community trends. Next, representative cases are selected for review, and permission is obtained from local law enforcement to review the cases. OFR leaders then coordinate the collection of available and pertinent information on the decedents across multiple sources, which may include their own sources, agency partners, and next-of-kin interviews with the decedent's family or friends. After compiling the data, they prepare a presentation for the case review meeting, often including a timeline of the decedent's interactions with various agencies. During the sessions, OFR members, agency partners, and community representatives review the data and collaboratively brainstorm strategies for overdose prevention.

Case Selection

Participants reported several factors that influence case selection. About 50% (n=5) of the participants reported looking at demographic trends such as age, race, and sex. About 50% (n=5) of the participants reported combining multiple demographic and substance trends into a single theme and looked explicitly at decedents within that theme. For example, 1 participant stated, "Now that we're doing theme selection, we may focus on a specific drug, like, I think the next theme that we're doing is African American men between certain ages that historically used cocaine, but fentanyl was also involved in their cause of death." Around 60% (n=6) of the participants prioritize selecting cases with comprehensive data, though they noted that this is challenging due to limited data availability during the initial selection process. Other key factors influencing case selection included obtaining permission from agency partners (n=5, 50%) and ensuring cases fell within jurisdictional boundaries (n=9, 90%).

Case Data Collection and Preparation

OFR leaders reported collecting data on decedents using publicly available databases and resources provided by their health departments and from the state-level data provided. They also reported requesting data from agency partners regarding any interactions with the decedent. After data are collected, they are processed manually by the OFR leader. One participant stated, "Then those individual partners have to go in, look at the specific case and then they have to like hand-put in all of the info and then they send those to me and then I scan them into our system so that we have them electronically, and then I have to take all of those electronic copies and upload those, one question at a time into REDCap." After processing, all participants reported compiling the information into a timeline to display during case review presentations, which helped viewers to understand the decedent's story.

Reported Challenges

When asked about challenges, 90% (n=9) of the participants identified a reliance on manual processes to collect data as a significant challenge. Due to limited bandwidths, responses to data requests from agency representatives are often delayed or incomplete, which impacts the preparation of case review materials. When final requests were not fulfilled, critical data were missing from presentations. Participants also reported several technological challenges, including siloed data sources (n=4, 40%), confusing data formatting (n=7, 70%), and other technological issues (n=10, 100%), all of which impacted data collection and preparation. As a result, 90% (n=9) of the participants highlighted time pressure as a major challenge for preparing for the OFR process. One participant explained this time pressure stating, "Lots of times for overdose fatality review, and this is true for me also, this is one part of my job, right? It's part of my FTE, it is not my full FTE, so there are directions that I'm pulled for other projects that can sometimes limit the time that I have available to work on this."

Desired Future State

Half (n=5) of the participants indicated that easier access to current data sources would be helpful, specifically mentioning simplified access as well as fewer lags to be able to identify and respond to current trends. One participant stated, "I think something that local public health and us specifically have always sort of struggled with is being able to keep up with that data…having it be a little easier for us to get local data and more quickly, that is an issue that is huge for us. We see things, and we hear about these trends, but we don't necessarily always know that that's happening until all of a sudden, it's like, hey, we're seeing this, you know, all across the county, and is that something we could have caught sooner had we been able to access that data quicker."

Participants mentioned that additional data would help them. About 70% (n=7) of the participants desired health care data, most commonly substance use disorder treatment data; however, these data are protected by federal statute, which adds complexity to accessing and sharing them. About 50% (n=5) of the participants desired criminal justice data, whether they did not have access to it or did not often receive it when requested. Participants mentioned that they worked with multiple law enforcement agencies and regularly received data from some but not others. Additionally, 2 participants mentioned that they had not been able to perform next-of-kin interviews due to barriers in setting up interviews. Other desired data sources included childhood information from school districts or Child Protective Services and input from local organizations.

About 80% (n=8) of the participants mentioned that more collaboration or support would help the OFR process. Other participants specifically mentioned that increased collaboration between health departments and local agencies would help their current workflow, create the possibility to expand their services, and address long-term sustainability of OFR.

Phase 2: Data Dashboard Design and Evaluation

Dashboard Design

The synthetic data dashboard prototype, developed based on Phase 1 data, was refined into a high-fidelity prototype for further usability and human factors evaluation. The final prototype featured 9 theme-based aggregate data pages covering demographics, substances used, health care interactions, prehospital emergency services, social and economic factors, mortality, prescription patterns, and treatment and recovery. Select components of the line-level data and timeline are shown in Figures 1 and 2, and a full demonstration video is available in our GitLab repository [23]. The final dashboard prototype included advanced filtering capabilities, enabling users to refine data by specific time frames, substances, and death status. A total of 3 machine learning tools were integrated into the dashboard. First, the Hospital Note Topics Tool utilized latent Dirichlet allocation for topic modeling of EHR notes [25]. This tool identified prevalent themes and trends, such as patterns of substance use, health care utilization, and social determinants of health, providing users with a high-level understanding of key insights from unstructured text data. Second, the 30-Day Risk Score for Readmission and Death employed an eXtreme Gradient Boost machine learning model, which analyzed a combination of EHR notes and tabular data along with EMS and neighborhood census data to predict the likelihood of hospital readmission or death within 30 days. Third, the EMS Geographic Prediction Tool combined EMS response data with neighborhood-level census tract information to identify geographic areas at higher risk for overdose events. Additionally, the final prototype incorporated a drill-down feature, allowing users to filter patient populations by category and narrow them down to individual patients.

Surveys

A total of 6 OFR organizers participated in the Phase 2 survey, representing stakeholders from county, state, and federal agencies. Additional demographic details are provided in Table 1. All the 6 participants reported analyzing aggregate data and selecting review cases, while 4 were involved in abstracting data for case reviews. The UTAUT results indicated that the dashboard would likely be adopted if made available to participants (Figure 4), with a mean of 4.07 out of 5.00 (SD 0.65). The Perceived Usefulness Questionnaire results suggested a moderately positive perception of usefulness for the aggregate and individual-level data (Figure 5), with means of 3.61 (SD 0.82) and 3.64 (SD 0.85) out of 5.00, respectively.

Figure 4. Unified Theory of Acceptance and Use of Technology (UTAUT) survey results for overdose fatality review leaders in Wisconsin, 2024.

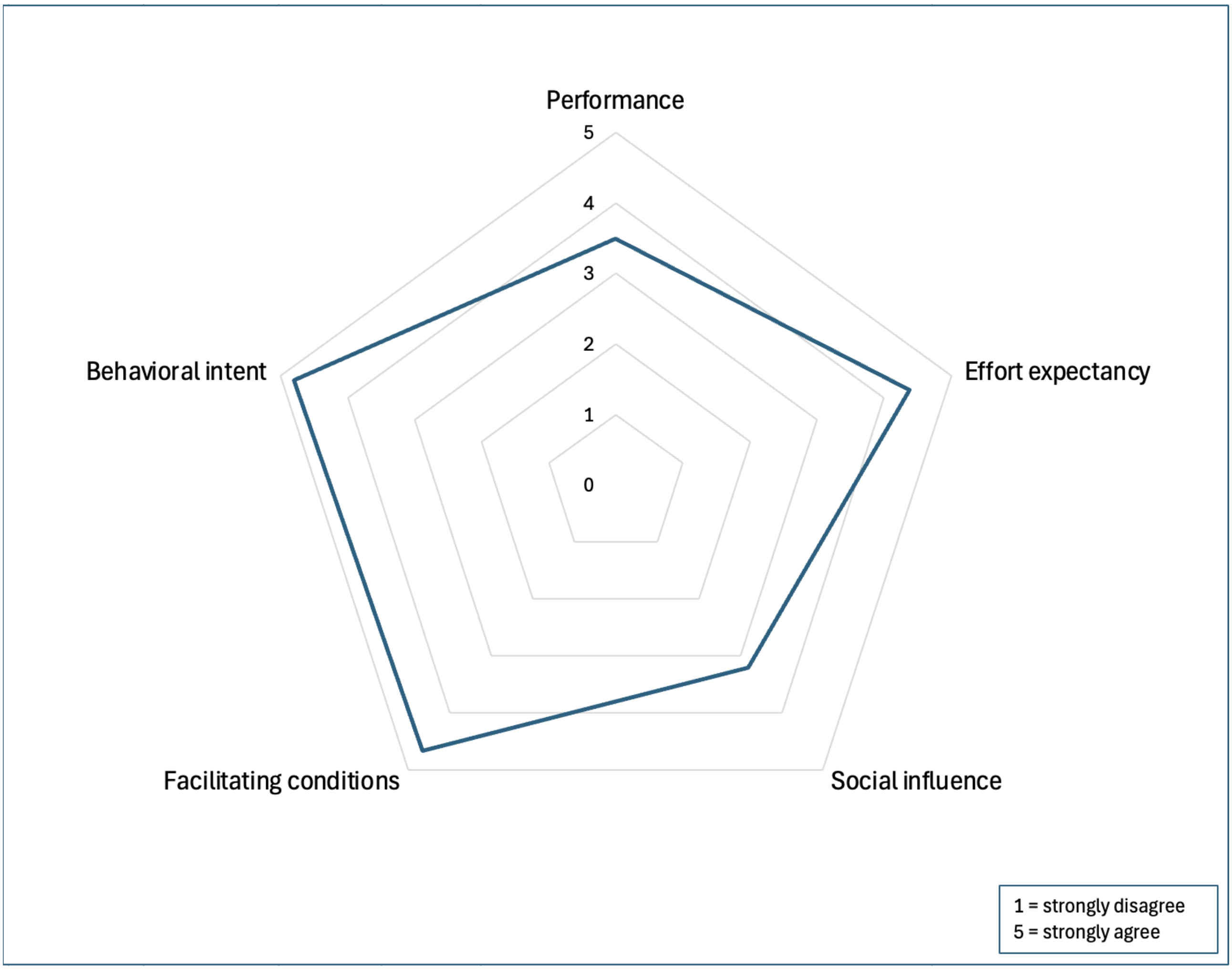
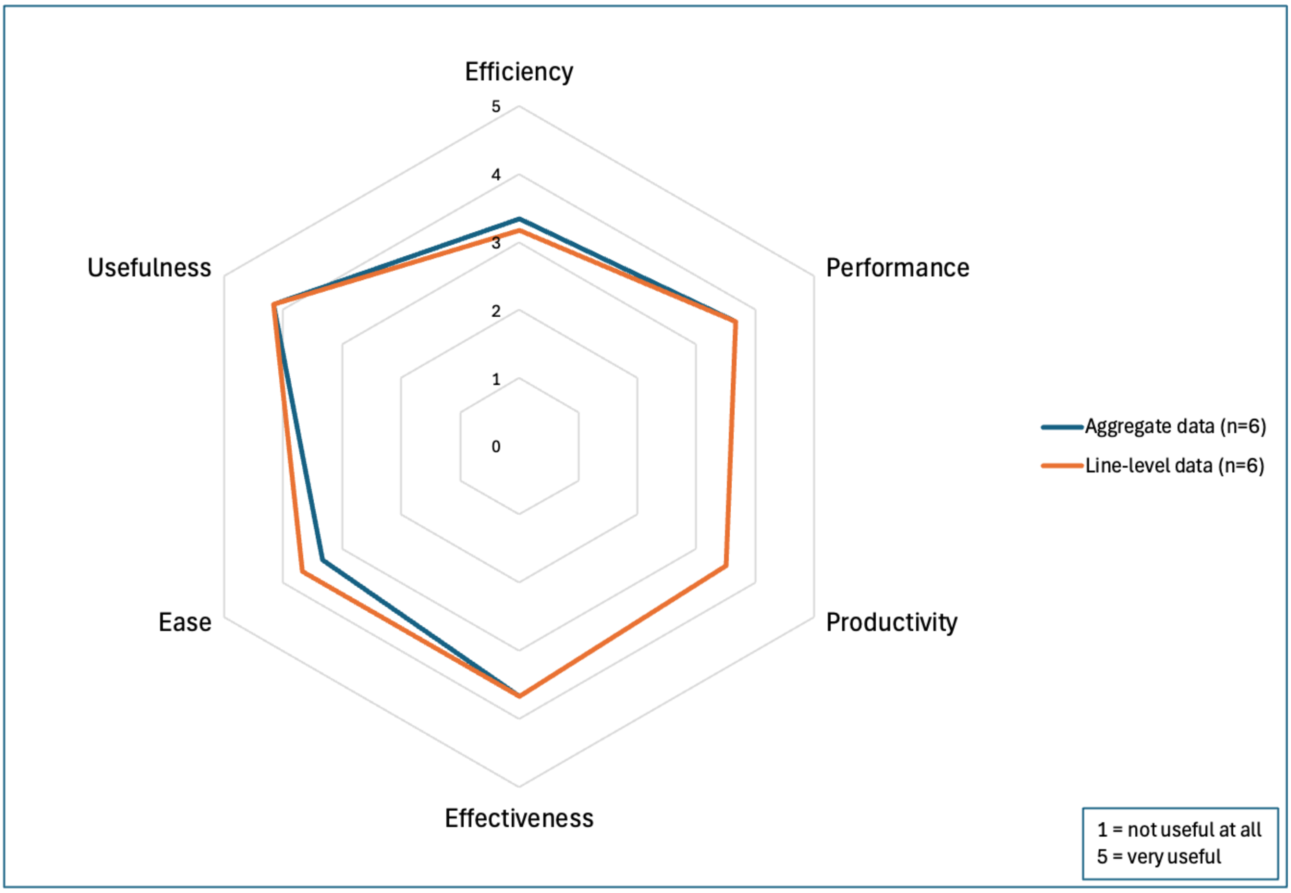


Figure 5. Perceived usefulness questionnaire results for overdose fatality review leaders in Wisconsin, 2024.



Interviews

A total of 7 OFR leaders participated in semistructured interviews. Participants were specifically asked about the potential benefits of the dashboard as well as areas for improvement. The summaries of the findings are outlined below, with detailed results and quotes provided in Multimedia Appendix 2.

Benefits of the Dashboard

All the participants (n=7) highlighted improved data access as a key benefit of the dashboard. This included expanding access to currently unavailable data sources, facilitating quicker and easier access to existing data, and increasing access for less-resourced communities. Additionally, 86% (n=6) of the participants indicated that the dashboard could help to optimize workflows by reducing the manual processes and introducing time-saving features. One participant stated, “That’s helpful in just kind of like streamlining it all in one spot, because the data that I do get like from emergency department visits, I have to go through every single entry, and then like, figure it out from there. So, this is quite nice to just have it in one spot where I can look at it, write it down, we’re done.” About 71% (n=5) of the participants noted that the dashboard’s tools are especially valuable for organizational outreach efforts. The machine learning tools, particularly the EMS geographic prediction tool and the 30-day risk score for readmission or death, were highlighted as most useful for these initiatives. All the participants (n=7) reported that the dashboard appeared easy to navigate, with an intuitive structure and organization. About 71% (n=5) of the participants specifically mentioned the ability to filter patients for specific populations as a standout feature, distinguishing the dashboard from other available tools.

Areas for Improvement

Most participants suggested adding more data sources. Specific recommendations included sexual orientation and gender identity, medical comorbidities such as chronic pain, and presenting data as rates instead of counts to better represent minority group trends. Concern about the accuracy and timeliness of the data was expressed by 71% (n=5) of the participants. The SMDC cohort, which is limited to patients with hospital encounters linked to other data, excludes individuals seen only by EMS or those without EMS or hospital contact. One participant expressed this concern, stating, “I like the way that it’s laid out, but because it’s only people who are going to the ED, I don’t know that we can draw conclusions about folks that are at risk of overdose, generally speaking.” The current data were also identified as a critical need for the dashboard. Participants proposed several technological enhancements to improve the dashboard’s usability. Specifically, 43% (n=3) of the participants suggested the ability to export data from the dashboard, while 57% (n=4) recommended adding other technological features such as additional filters and hover-over tips for ease of use. Other important but less commonly cited concerns included a training or learning curve (n=3, 43%), having jurisdictional access (n=2, 29%), adding to stigma or bias (n=2, 29%), having too many years aggregated in the dataset to be reflective of current trends (n=2, 29%), and the dashboard data being deidentified and therefore not being able to be connected to data about individuals from other sources (n=2, 29%).

Discussion

Principal Findings

This study is the first to systematically assess the cognitive workload of OFR leaders and apply those findings to the design of a user-centered data dashboard prototype. Using validated instruments (NASA-TLX, UTAUT, and Perceived Usefulness Questionnaire) and qualitative interviews, we identified substantial cognitive demands associated with data aggregation, case selection, and abstraction across multiple agencies. Guided by these findings, we developed and tested a prototype dashboard using synthetic data modeled on the SMDC to streamline data access and visualization. Participants reported improved usability, reduced time burden, and high potential for adoption in future OFR workflows.

Interpretation and Implications

Our results highlight that OFR leaders face high mental and temporal demands due to fragmented data sources and reliance on manual processes, a challenge consistent with prior studies on multiagency data integration and decision-making in public health surveillance [26]. Previous dashboards have been implemented at the state or regional level to visualize aggregate overdose trends, yet few, if any, have incorporated both population- and individual-level data or been explicitly designed around the OFR workflow. By linking health system, emergency response, and correctional datasets within a privacy-preserving synthetic environment, our approach extends beyond traditional epidemiologic dashboards to directly support operational review teams. The integration of predictive modeling tools, such as 30-day readmission and EMS geographic risk scoring, further distinguishes this work from prior descriptive dashboards by introducing prospective, data-driven decision support [7,8]. These findings align with recent public health informatics initiatives emphasizing human-centered design, cognitive workload reduction, and automation to improve decision-making efficiency [27,28].

Our SMDC data dashboard serves as a comprehensive tool for extracting, transforming, and visualizing overdose data. It was designed to refresh with recent case information from multiple sources; incorporate automation for case matching, data formatting, and quality checks; and offer easy navigation to streamline current workflows. The dashboard helps automate the current workflow challenge of siloed data sources and includes additional data that participants found valuable. It integrates key data variables from multiple agencies, with options to filter the data by important demographic and substance-related factors. The automated timeline feature compiles data from all sources in our dataset, visually representing the events leading to an overdose death. Phase 2 participants saw the dashboard's potential to reduce time pressure and reliance on manual processes. Additionally, in larger cities where reviewing every case is impractical or impossible, OFR leaders typically hand-pick cases to analyze in detail. By linking all cases and presenting the data in aggregate, this system enhances the understanding of overall trends and helps to better inform recommendations.

Many participants indicated that they would use this dashboard as an additional tool, rather than replacing their current methods. This likely explains our Phase 2 survey findings, which revealed a discrepancy between the tool's effectiveness and efficiency. This preference highlights a key concern raised in Phase 2 interviews that cannot be resolved in future dashboard versions: the deidentified nature of the SMDC data prevents integration with other data sources, such as next-of-kin interviews, a crucial part of the case review process. Other areas of improvement discussed in Phase 2 interviews can be resolved.

Phase 1 participants emphasized a desire for increased collaboration and support. Some mentioned this in the form of improved relationships between the health department and other agencies. Expanding the sectors involved in OFR processes and broadening the information available for case reviews may highlight previously unseen gaps in care. Emergency department utilization is common among those who misuse opioids and other drugs, and the number of emergency department visits is associated with an increased risk of drug overdose [29]. Our dashboard aims to establish and improve data sharing between OFR teams and health care systems, which has been identified as an important prevention strategy implementation [30].

Our findings underscore the potential of informatics-driven tools to enhance collaboration among public health, health care, and community agencies. By reducing cognitive workload and manual data handling, such tools can accelerate case review preparation, standardize data access across jurisdictions, and allow OFR teams to focus more on interpretation and prevention strategies. The use of synthetic data generated through large language models also represents an innovative method for tool development when working with sensitive, multiagency datasets. As counties nationwide expand their OFR infrastructure, the approach demonstrated here offers a transferable model for building and evaluating data dashboards that are secure, scalable, and adaptable to local contexts.

Limitations

The general process reported by participants in the interviews is mainly consistent with the process in the OFR Practitioner's Guide and the Public Health and Safety Team toolkit, which are guiding frameworks for health departments when creating OFRs and holding case review meetings [31]. However, the feedback in this study was collected from local OFR leaders in Wisconsin, and therefore, these findings may not be generalizable to other systems. Limitations inherent to the SMDC include its inclusion criteria and its deidentified nature. All data variables included in the prototype dashboard are included in the SMDC; however, the SMDC does not include all data variables important to OFR leaders when presenting cases. Finally, addressing challenges in the OFR process with new technology may not ultimately lead to better outcomes. Substance misuse is both a complex medical condition and an evolving public health issue. In order to make meaningful progress and improve outcomes, sustained collaboration across health care systems, public health agencies, and communities will be essential to reform policies, reduce disparities, and improve medical care.

The existing OFR process is built on a thorough, team-based approach, but it includes several cognitively demanding tasks, and there are multiple challenges to timely data preparation. Increased collaboration, access to standard, centralized tools, and comprehensive data could build upon the rigorous work already being done by OFR teams in order to further augment and automate workflows to reduce manual work. We designed a user-centered data dashboard to help reduce the cognitive workloads identified from surveys and incorporate the desired data sources and workflows gathered from the interviews. Evaluative feedback indicated many potential benefits as well as some areas for improvement. This insight will guide the development of a real-time data dashboard accessible to OFR leaders in their review process.

Conclusions and Broader Implications

Despite these limitations, this study demonstrates a replicable, human-centered approach for modernizing OFR workflows through data integration and visualization. The combination of cognitive workload analysis, synthetic data modeling, and dashboard-based decision support provides a foundation for scalable OFR modernization across states. Beyond overdose prevention, this framework illustrates how human factors and data science methods can be combined to enhance other multidisciplinary public health review processes. Future work will focus on deploying the dashboard with live SMDC data, expanding data sources, and assessing real-world impacts on timeliness, data completeness, and prevention outcomes.

Funding

This work was supported by the National Institutes of Health, National Institute on Drug Abuse (R01DA051464), and the Department of Medicine of the University of Wisconsin School of Medicine and Public Health (233-AAM7146). The content is solely the responsibility of the authors and does not necessarily represent the official views of the UW-Madison Department of Medicine.

Data Availability

The full set of National Aeronautics and Space Administration-Task Load Index, Unified Theory of Acceptance and Use of Technology, and Perceived Usefulness Questionnaire questions; our interview guides and codebooks; our synthetic dataset and prompts; and our demonstration video are all viewable in our GitLab repository [23].

Authors' Contributions

Conceptualization: MKO, MP, AM, MA
Data curation: AM, MA, MKO, MP
Formal analysis: MP, AM, AP, CK, JDG, MA, MKO
Funding acquisition: AM, MA, MKO
Investigation: MKO, MP
Methodology: MP, MKO, JDG, CK, AP, ESA, BP, MS, MG, AM, MA
Project administration: AM, MA, MKO, MP
Resources: AM, AP, BP, CK, ESA, JDG, MA, MS, MG, MKO
Software: AK, AM, AP, ASA, CK, JDG, MA, MKO, PG, TG
Supervision: AM, MA, MKO
Validation: AM, AP, CK, ESA, JDG, MA, MS, MG, BP
Visualization: AK, AM, ASA, BP, MA, MG, MKO, MS, MP, PG, TG
Writing – original draft: AM, MA, MKO, MP
Writing – review & editing: MP, MKO, JDG, CK, AP, SHM, ESA, BP, MS, MG, AK, TG, PG, ASA, AM, MA

Conflicts of Interest

None declared.

Multimedia Appendix 1
Phase 1 interview results for overdose fatality review leaders in Wisconsin, 2023-2024.
[DOCX File, 21 KB - humanfactors_v13i1e79407_app1.docx]

Multimedia Appendix 2
Phase 2 interview results for overdose fatality review leaders in Wisconsin, 2024.
[DOCX File, 21 KB - humanfactors_v13i1e79407_app2.docx]

Checklist 1

COREQ checklist.
[PDF File, 378 KB - humanfactors_v13i1e79407_app3.pdf]

References

1. Spencer MR, Garnett MF, Miniño AM. Drug overdose deaths in the United States, 2002–2022. : National Center for Health Statistics; 2023 URL: https://stacks.cdc.gov/view/cdc/135849 [accessed 2024-06-20]
2. Robinson A, Christensen A, Bacon S. From the CDC: the prevention for states program: preventing opioid overdose through evidence-based intervention and innovation. J Safety Res 2019 Feb;68:231-237. [doi: 10.1016/j.jsr.2018.10.011] [Medline: 30876516]
3. 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;20(4):553-564. [doi: 10.1177/1524839918797617] [Medline: 30205737]
4. Afshar M, Oguss M, Callaci TA, et al. Creation of a data commons for substance misuse related health research through privacy-preserving patient record linkage between hospitals and state agencies. JAMIA Open 2023 Dec;6(4):ooad092. [doi: 10.1093/jamiaopen/ooad092] [Medline: 37942470]
5. Schulze A, Brand F, Geppert J, Böl GF. Digital dashboards visualizing public health data: a systematic review. Front Public Health 2023;11:999958. [doi: 10.3389/fpubh.2023.999958] [Medline: 37213621]
6. 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. [doi: 10.2196/humanfactors.9328] [Medline: 29853440]
7. Salvi A, Gillenwater LA, Cockrum BP, et al. Development of a real-time dashboard for overdose touchpoints: user-centered design approach. JMIR Hum Factors 2024 Jun 11;11:e57239. [doi: 10.2196/57239] [Medline: 38861717]
8. Fliss MD, Cox ME, Proescholdbell S, Patel A, Smith M. Tying overdose data to action: North Carolina’s opioid and substance use action plan data dashboard. J Public Health Manag Pract 2023;29(6):831-834. [doi: 10.1097/PHH.0000000000001796] [Medline: 37498535]
9. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
10. 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. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
11. Hart SG. NASA-task load index (NASA-TLX); 20 years later. Proc Hum Factors Ergon Soc Annu Meet 2006 Oct;50(9):904-908. [doi: 10.1177/154193120605000909]
12. Kallio H, Pietilä AM, Johnson M, Kangasniemi M. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. J Adv Nurs 2016 Dec;72(12):2954-2965. [doi: 10.1111/jan.13031] [Medline: 27221824]
13. Hewitt-Taylor J. Use of constant comparative analysis in qualitative research. Nurs Stand 2001;15(42):39-42. [doi: 10.7748/ns2001.07.15.42.39.c3052] [Medline: 12212430]
14. O’Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020 Jan 1;19:1609406919899220. [doi: 10.1177/1609406919899220]
15. Pollack AH, Simon TD, Snyder J, Pratt W. Creating synthetic patient data to support the design and evaluation of novel health information technology. J Biomed Inform 2019 Jul;95:103201. [doi: 10.1016/j.jbi.2019.103201] [Medline: 31078659]
16. GPT-4 [technical report]. : OpenAI; 2023 URL: https://openai.com/research/gpt-4 [accessed 2025-12-10]
17. Wei J, Bosma M, Zhao VY, et al. Finetuned language models are zero-shot learners. arXiv. Preprint posted online on Sep 3, 2021. [doi: 10.48550/arXiv.2109.01652]
18. YData synthetic data for machine learning. YData. 2024. URL: https://ydata.ai/ [accessed 2024-06-19]
19. Power BI. Microsoft Corporation. 2024. URL: https://powerbi.microsoft.com/ [accessed 2025-12-08]
20. Dowding D, Merrill JA. The development of heuristics for evaluation of dashboard visualizations. Appl Clin Inform 2018 Jul;9(3):511-518. [doi: 10.1055/s-0038-1666842] [Medline: 29998455]
21. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view1. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
22. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
23. Pisani M, Oguss M, Mayampurath A, et al. Substance misuse data commons dashboard. GitLab. 2024. URL: https://git.doit.wisc.edu/smph-public/dom/uw-icu-data-science-lab-public/substance-misuse-data-commons-dashboard [accessed 2025-12-10]
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;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
25. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993-1022. [doi: 10.5555/944919.944937]

26. Iyamu I, Gómez-Ramírez O, Xu AX, et al. Challenges in the development of digital public health interventions and mapped solutions: findings from a scoping review. Digit Health 2022;8:20552076221102255. [doi: 10.1177/20552076221102255] [Medline: 35656283]
27. Ansari B, Martin EG. Integrating human-centered design in public health data dashboards: lessons from the development of a data dashboard of sexually transmitted infections in New York State. J Am Med Inform Assoc 2024 Jan 18;31(2):298-305. [doi: 10.1093/jamia/ocad102] [Medline: 37330670]
28. Fareed N, Olvera RG, Wang Y, et al. Lessons learned from developing dashboards to support decision-making for community opioid response by community stakeholders: mixed methods and multisite study. JMIR Hum Factors 2024 Sep 9;11:e51525. [doi: 10.2196/51525] [Medline: 39250216]
29. 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):613-619.e2. [doi: 10.1016/j.annepidem.2015.03.018] [Medline: 25935710]
30. Wong CA, Houry D, Cohen MK. Integrating public health and health care-protecting health as a team sport. N Engl J Med 2024 May 16;390(19):1739-1742. [doi: 10.1056/NEJMp2403274] [Medline: 38598818]
31. Rubel S, Roe S. Public health and safety team toolkit: guidance for data-driven overdose response coordination among public health, criminal justice, law enforcement, and first responders. : CDC Foundation, Centers for Disease Control and Prevention; 2022 URL: https://stacks.cdc.gov/view/cdc/127746/cdc_127746_DS1.pdf [accessed 2025-12-18]

ABBREVIATIONS

COREQ: Consolidated Criteria for Reporting Qualitative Research
EHR: electronic health record
EMS: emergency medical services
ICR: intercoder reliability
NASA-TLX: National Aeronautics and Space Administration-Task Load Index
OFR: overdose fatality review
REDCap: Research Electronic Data Capture
SMDC: Substance Misuse Data Commons
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 14.Jul.2025; peer-reviewed by AL Dir, K Schwartz; revised version received 07.Oct.2025; accepted 12.Nov.2025; published 15.Jan.2026.

Please cite as:
Pisani M, Oguss MK, Dickson-Gomez J, Kostelac C, Parry A, Moss S, Salisbury-Afshar E, Patterson B, Spigner M, Gussick M, Krautkramer A, Gruenloh T, Safipour Afshar A, Gupta P, Mayampurath A, Afshar M
Designing a Substance Misuse Data Dashboard for Overdose Fatality Review Teams: User-Centered Design Approach
JMIR Hum Factors 2026;13:e79407
URL: https://humanfactors.jmir.org/2026/1/e79407
doi:10.2196/79407

Original Paper

Navigation, Adoption, and Use of Digital Health Technologies for Irritable Bowel Syndrome Self-Management: Focus Group Study of Patient Experience and Decision-Making

Adrijana D'Silva[1*], PhD; Nicolle Hua[1*], MGenCouns; Mary V Modayil[2], PhD; Judy Seidel[1,2], PhD; Deborah A Marshall[1,3,4], PhD

[1]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[2]Primary Care Alberta, Edmonton, AB, Canada

[3]Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[4]O'Brien Institute for Public Health, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[*]these authors contributed equally

Corresponding Author:
Deborah A Marshall, PhD
Department of Community Health Sciences
Cumming School of Medicine
University of Calgary
HRIC 3C60
3280 Hospital Drive NW
Calgary, AB, T2N 4Z6
Canada
Phone: 1 403 210 6377
Email: damarsha@ucalgary.ca

Abstract

Background: Irritable bowel syndrome (IBS) is a common chronic gastrointestinal disorder that impairs bowel functions and patients' overall quality of life. IBS-focused digital health technologies (DHTs), including online health resources and mobile health (mHealth) apps, have recently proliferated for patient use. However, research exploring patients' experiences with navigating, adopting, or using commercial or publicly available DHTs for IBS self-management is limited.

Objective: This study aims to explore the user experiences and decision-making of patients with IBS as they navigate, adopt, and use diverse DHTs for disease self-management.

Methods: We conducted virtual semistructured focus group interviews to explore the experiences of patients with IBS using DHTs, including their perspectives on design and features, their decision-making process in using DHTs, and recommendations for improving user experience and uptake, given the heterogeneous nature of these tools. Canada-based patients with IBS who were using or had used mHealth apps to manage symptoms were recruited through purposive sampling from previous IBS-related studies. Discussions were transcribed verbatim, and inductive thematic analysis was performed using NVivo (version 14; Lumivero). A modified version of the Expanded Unified Theory of Acceptance and Use of Technology (UTAUT2) model was applied to guide the interpretation of the dynamic relationship between the influences on participants' decisions regarding DHT use.

Results: Among the 8 participants (all female; mean age 55.3, SD 13.5 years), two themes were identified: (1) uncertainty impacts the trustworthiness of DHTs, and (2) influences that drive the decision-making process to adopt and use DHTs. The observed influences aligned with the constructs of the UTAUT2 model (performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit), with the addition of trust and risk in participants' decision-making. Digital health literacy and patient engagement were also raised as crucial components of participants' experiences and perspectives on DHTs.

Conclusions: Findings of this study highlight the current landscape of digital health in IBS and existing gaps and challenges for patients in navigating, adopting, and using DHTs for IBS self-management. While DHTs were generally viewed positively for their value and potential, patients with IBS consider several coexisting factors and trade-offs in their decision-making. Further investigations on the influences on and perspectives toward DHTs could enhance future development and iterations of these tools and improve patient confidence and uptake.

(JMIR Hum Factors 2026;13:e75012) doi:10.2196/75012

KEYWORDS

decision-making; digital health; health literacy; irritable bowel syndrome; mobile applications; patient experiences; patient participation; qualitative study; self-management; technology

Introduction

Digital health has revolutionized health care and disease management by leveraging technology and digital devices to improve self-management. With a growing public focus on health and well-being, more patients with chronic diseases rely on digital health, especially online health resources and mobile health (mHealth) apps [1,2], for answers, guidance, and management support.

Irritable bowel syndrome (IBS) is a disorder of the brain–gut connection that impacts patients' daily activities and well-being and is characterized by abdominal pain, bloating, altered bowel habits, and psychiatric comorbidities [3-5]. As there is no known cure for IBS, greater knowledge of the syndrome and relevant self-management options are critical in alleviating symptoms and improving quality of life [6,7]. However, IBS self-management strategies—such as lifestyle and dietary modifications and brain–gut behavior therapies [8,9]—are often coupled with challenges around accessibility, adherence, affordability, and productivity loss for patients [10-12]. Digitizing health information and support could offer alternative models of delivery and potentially address these issues by improving access to information and treatments, reducing health-related costs, and enabling and empowering patients [13-16].

In the context of IBS, digital health technologies (DHTs) encompass a diverse range of tools that support dietary, psychological, behavioral, and educational aspects of self-management. Commonly used dietary-focused tools include low fermentable oligosaccharides, disaccharides, monosaccharides, and polyols (FODMAP) diet apps, which are supported by evidence for improving symptom control and dietary adherence [17]. Psychological and brain–gut behavior apps have also demonstrated effectiveness in reducing IBS symptom severity and improving quality of life [18,19]. In addition, patients frequently use general symptom trackers, medication apps, and online educational resources of varying quality. This heterogeneity underscores the importance of understanding not only whether patients use DHTs, but also the specific domains these tools address, as each represents a difference in therapeutic mechanisms and levels of evidence within IBS care.

Our understanding of DHTs for IBS remains limited. Although several studies support the benefits of DHTs for patients with IBS [20], there is limited evidence on their experiences in a consumer context. Existing studies that qualitatively evaluate DHTs, although informative, have been limited to specific DHTs under controlled conditions [21-23] and otherwise lacked the sociodemographic context that could otherwise impact uptake and use. Other studies suggest acceptance of DHTs is determined by users' hedonic motivations, user-system fit, and perceived utility and intuitiveness of the technology [24-26]. However, it is uncertain how these influences or other priorities influence this specific patient cohort. Currently, it is difficult to assess other considerations and consequences for the real-world application of DHTs for the people with IBS. This study is an initial investigation to explore the experiences of patients with IBS in navigating, adopting, and using online health resources and mHealth apps to support self-management, and to identify key influences in their decisions to consider DHTs.

Methods

Ethical Considerations

This study was reviewed and approved by the University of Calgary Conjoint Health Research Ethics Board (REB23-1273). A qualitative study design with semistructured focus groups was used to explore the experiences and decision-making processes of patients with IBS regarding DHTs, including their adoption, use, and suggestions for improving these tools. The study methodology was reported in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Multimedia Appendix 1) [27].

Positionality

The study team comprised 5 women with research experiences in health systems (AD, NH, MVM, JS, and DAM), including one (AD) with lived experiences of IBS. AD and NH, a postdoctoral fellow and research associate, respectively, were skilled in qualitative methods. AD and DAM, a health economist and health services researcher, also had experience in IBS-related research. MVM and JS, a senior scientist and scientific director, respectively, had applied research expertise and experience in primary care. All but one each have a PhD; NH has a master's degree. Information on AD's personal experience and professional background in IBS was shared with study participants prior to the focus group sessions. Some participants had an existing relationship with AD from previous research studies.

Participants

Participants were recruited through purposive sampling from previous IBS-related studies conducted by the study team. Past participants who had provided consent to be recontacted for future studies were invited by email. Convenience sampling was attempted using the study team's social media networks and the University of Calgary's research recruitment board;

while 2 individuals were recruited through this method, neither participated in the focus groups. All participants were adults with IBS living in Alberta and British Columbia, Canada, who (1) had been formally diagnosed by a health care professional and (2) had been using or had used mHealth apps to manage their symptoms. Participants who satisfied the selection criteria received an information document regarding the study and provided written consent before the focus group sessions.

The study team a priori aimed for at least 4 focus groups, with 3 participants per group, which was expected to be sufficient to reach code saturation [28-30]. While focus group size typically ranges from 6 to 8 participants, smaller groups with 3 participants were considered to allow more opportunities for engagement and account for the potentially low recruitment yield [28]. The recruitment process ended following the exhaustion of all recruitment strategies, resulting in 8 participants and 3 focus groups. Some degree of saturation was present in participants’ responses through repeated codes and themes.

Data Collection

Participants completed a survey before attending the focus group session to provide details of their existing supports, demographic information, year of IBS diagnosis, symptoms, treatment history, and dietary adjustments (Multimedia Appendix 2). The survey and data were managed using Qualtrics Experience Management (Qualtrics LLC) and the University of Calgary’s secured online cloud server. Survey data were analyzed and summarized using Microsoft Excel by Darren Zhang.

Focus groups were conducted virtually using Microsoft Teams. AD moderated the focus group discussions, with the support of NH as a scribe. Sessions lasted approximately 2 hours, were audio-recorded using the recording feature in Microsoft Teams, and were transcribed verbatim using the transcription service Rev (Rev.com, Inc); the transcripts were subsequently reviewed and cleaned by Darren Zhang.

Instrumentation

Participants’ IBS symptoms were assessed in the survey using the IBS Symptom Severity Score (IBS-SSS); scores ranged from 0 to 500 and were categorized as mild (75-175), moderate (176-300), or severe (>300) [31]. A focus group guide was created and reviewed by the study team to support AD in facilitating semistructured sessions (Multimedia Appendix 3). The following topics were discussed in the focus groups: (1) participants’ experiences navigating digital resources to self-manage IBS, (2) their experiences adopting and using DHTs to self-manage IBS, and (3) their perspectives on influences and improvements regarding the uptake of online resources and mHealth apps.

Data Analysis

Coding was performed and managed using NVivo (version 14; Lumivero). NH and AD coded the transcripts as primary and secondary coders, respectively. Because little research has explored this topic, inductive thematic analysis was used to identify and evaluate relationships, patterns, or themes, allowing data to serve as the foundation of the results [32]. Initial codes were generated to indicate recurring features within participants’ responses, and related codes were then collated into groups that subsequently defined underlying themes [32]. Following thematic analysis, the findings and their relationships were visualized based on the Expanded Unified Theory of Acceptance and Use of Technology (UTAUT2) framework model [26].

Applied Framework

The UTAUT2 framework model is an extended version of the original UTAUT model to evaluate the acceptance of new commercial technologies [26,33]. With its consumer-use context, this model was chosen to provide valuable and holistic insight into patients’ acceptance of DHTs, including adoption and use, for IBS self-management. The key constructs of the UTAUT2 model [26] included performance expectancy, defined as the extent to which the individual expected the performance or benefits of the technology; effort expectancy, defined as the level of effort or ease in using the technology by the individual; social influence, defined as the extent to which the individual perceived the value of others’ acceptance of the technology; facilitating conditions, defined as the extent to which the individual believed that support or resources were available to facilitate the use of the technology; hedonistic motivation, defined as the extent of the individual’s motivations or pleasures experienced in using the technology; price value, defined as the cost-benefit or trade-off between the perceived value or benefit of the technology and its associated monetary cost; and habit, defined as the extent to which the individual developed a habitual or perpetual use of the technology from previous experience and learning.

Also, we modified the UTAUT2 model to include “trust” and “risk” as key constructs based on our thematic findings. Although they were not included in the original UTAUT2 model [26], previous studies have recognized and incorporated them into their own models [34-36]. Based on these studies [34-36], “trust” was defined as the extent to which an individual perceived the technology or its associated qualities as likable or trustworthy, and their willingness to believe that the technology would meet their expectations, and “risk” was defined as the extent to which an individual perceived the technology to be associated with potential consequences or harm to themselves.

Additionally, DHT-related key constructs are influenced by user characteristics. In the expanded UTAUT2 model, Venkatesh et al [26] recognized how a user’s demographic characteristics, specifically age, gender, and experiences or familiarity of existing technologies, moderate the impacts of the key constructs and ultimately the acceptance and use of the technology.

Results

Participant Characteristics

Of the 11 participants recruited to the study, 8 attended the focus groups, all of whom had participated in past research studies. One individual was lost to follow-up prior to the focus group sessions, and the remaining 2 could not attend the sessions due to scheduling conflicts. Participant characteristics are summarized in Table 1.

Table 1. Participants' characteristics (N=8).

Characteristics	Value
Age (years), mean (SD)	55.3 (13.5)
Gender (female), n (%)	8 (100)
Highest education, n (%)	
Postsecondary certificate or diploma below bachelor level or apprenticeship or trades certificate	1 (13)
Bachelor's degree	1 (13)
Postsecondary certificate or diploma above bachelor level	4 (50)
Master's degree	2 (25)
Identify as, n (%)	
Born outside of Canada	3 (38)
Persons with disabilities	2 (25)
2SLGBTQI+[a]	1 (13)
Racialized or visible minority	1 (13)
Indigenous	0 (0)
I have children or grandchildren aged 18 years or younger living at home	0 (0)
New to Canada (less than 5 years)	0 (0)
Other	1 (13)
None of the above	3 (38)
Years since IBS[b] diagnosis, mean (SD)	12.9 (10.6)
Types of IBS, n (%)	
Mixed	2 (25)
Constipation-dominant	5 (63)
Diarrhea-dominant	1 (13)
Unsure	0 (0)
IBS-SSS[c], mean (SD)	197.7 (84.8)
Types of digital tool used, n (%)	
Mobile health apps	7 (88)
IBS-related podcasts	4 (50)
Mobile tracking apps (eg, physical activity and daily living activities)	6 (75)
Websites	6 (75)
Support groups	1 (13)
Types of mobile health apps (n=7), n (%)	
Diet-related apps	2 (29)
IBS-specific apps	6 (86)
Mental health or emotional well-being apps	1 (14)
General gastrointestinal and health apps	2 (29)
Cognitive-behavioral therapy and/or gut hypnotherapy apps	1 (14)

[a] 2SLGBTQI+: two-spirited, lesbian, gay, bisexual, transgender, queer/questioning, intersex, and others.

[b] IBS: irritable bowel syndrome.

[c] IBS-SSS: IBS Symptom Severity Score; ranges from 0 to 500, with mild (75-174), moderate (175-299), and severe (300-500) symptom severity.

All participants identified as women and had tertiary-level education. The mean age was 55.3 (SD 13.5) years, and participants had lived with an IBS diagnosis for an average of 12.9 (SD 10.6) years. Five of 8 (63%) participants reported constipation-dominant IBS, 2 (25%) reported mixed IBS, and 1 (13%) reported diarrhea-dominant IBS. The average IBS-SSS was 197.9 (SD 84.8), indicating mild symptoms. Participants had managed their IBS symptoms using one or a combination of the following DHTs: a total of 7 (88%) had used mHealth apps, 6 (75%) used mobile tracking apps, 6 (75%) used websites, 4 (50%) used IBS-related podcasts, and 1 (13%) used support groups. Among all types of mHealth apps used (n=7), 6 participants (86%) reported having used IBS-specific apps, 2 (29%) diet-specific apps, 2 (29%) general gastrointestinal and health apps, 1 (14%) mental health and/or emotional well-being apps, and 1 (14%) cognitive-behavioral therapy and/or gut hypnotherapy apps.

Themes of Navigating, Adopting, and Using Digital Health Technologies

Two themes were identified from participants' experiences and decision-making around navigating, adopting, and using DHTs for IBS self-management: (1) uncertainty impacts the trustworthiness of DHTs, and (2) underlying influences drive the decision-making process to adopt and use DHTs.

Theme 1: Uncertainty Impacts the Trustworthiness of Digital Health Technologies

A sense of uncertainty was prevalent among participants when navigating, adopting, and using DHTs for their symptoms of IBS. While participants were generally positive toward DHTs with hope and expectations of effective self-management and symptomatic relief, the coexisting challenges and risks contributed to varying levels of distrust.

Participants expressed uncertainty toward the effectiveness, utility, and understanding of DHTs. Because of the self-directed nature of the navigation process through their mobile devices and the internet, participants were aware that not all DHTs would be effective or suitable for them, and some recounted their experiences of finding suitable DHTs through trial-and-error.

> *[. . .] As we know, it is hard to know if anything's working. It's hard to know if anything works for the symptoms that we're experiencing. So that does come into it as well, but it's not a science, it's sort of a bit of a feeling about whether you're going to continue with something and whether it's worth it.* [Participant #48]

> *I just found and tried and listened to whatever. And then if it wasn't in my wheelhouse then I just discarded.* [Participant #29]

In contrast to those with positive experiences, other participants with subpar experiences using DHTs expressed challenges and frustration in finding what they needed to manage IBS. The lack of guidance around appropriate and trustworthy DHTs left some feeling alone and overwhelmed by the amount of information online or the number of mHealth apps to consider.

> *We just kind of have to search and go online and try to look at reliable resources [. . .] I've got family help too that they're in the tech side of things, so I get some help from them from looking at different options. But basically, you do it on your own.* [Participant #37]

> *You're feeling crummy. You need a resource, you need some help, but trying to wade through everything or do a Google (search) is just so overwhelming.* [Participant #9]

Uncertainty toward the DHTs' intentions was raised, including fear of scams or malicious attacks from illegitimate websites and privacy and data exploitation from mHealth apps. Furthermore, participants discussed the accuracy of online information from the internet and their concerns about consuming IBS-related misinformation and disinformation.

> *I agree with (Participant #9) regarding credentials and if (doctors are) trying to sell stuff. I know there's some doctors out there who do that too. So yeah, just mindful of people trying to market products versus people with credentials who are trying to debunk misinformation or provide just evidence backed information.* [Participant #14]

> *I don't like even giving my credit card information or any other personal information out into the internet. You just never know who's out there and what's going to happen to it.* [Participant #37]

During these discussions, digital health literacy was raised as an essential skill and a common barrier for patients with IBS to assess, process, and understand digital health information and technologies.

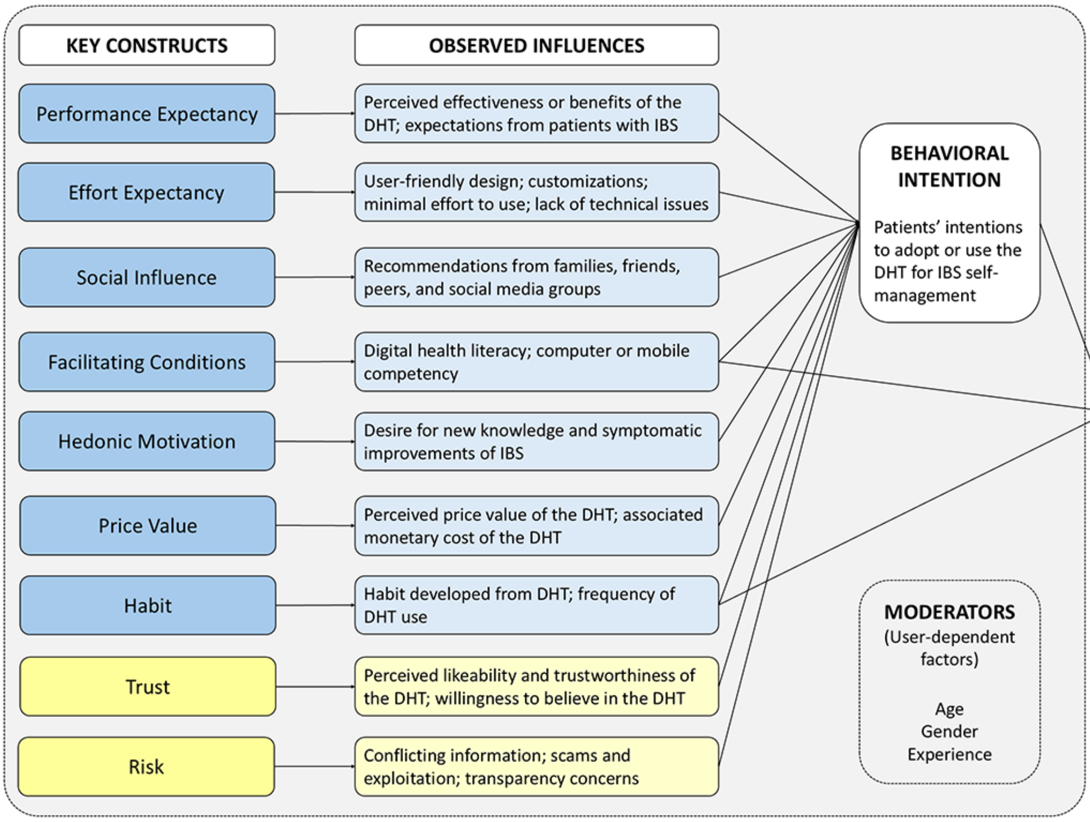
> *It's not about (like) or dislike, it's always questioning what I found, is it true or not, and how to explore further to be sure that these things are true because I'm not a great researcher [. . .] because of lack of medical background. So I am always hesitant to trust this or that information. It's not disliking, but it's a concern whether it's right direction to move or not.* [Participant #4]

Theme 2: Influences That Drive the Decision-Making Process to Adopt and Use Digital Health Technologies

Overview

Participants alluded to several influences that contributed to their decision to use DHTs, which were contextualized using the modified UTAUT2 model (Figure 1). Applying the model to this cohort illustrates the model's relevance and the complexities of DHT acceptance among patients with IBS. The observed influences (key constructs), as proponents or opponents to their decision-making process at varying degrees, collectively impact participants' intentions and/or the actual adoption or use of DHTs. Furthermore, the influences may have been externally moderated by participants' demographics, including age, gender, and experience or familiarity with DHTs. Selected quotes are provided to illustrate the influences and their relation to key constructs from the UTAUT2 model (Multimedia Appendix 4).

Figure 1. Relationships of identified influences in the decision-making process of patients with irritable bowel syndrome (IBS) on the use and acceptance of digital health technologies (DHTs) based on the modified Expanded Unified Theory of the Use and Acceptance of Technology (UTAUT2) model. "Trust" and "risk" were not part of the original UTAUT2 model and were included as a result of the study's findings.



Performance Expectancy

Based on their prior knowledge, participants expressed varying levels of preconceived expectations toward the DHTs' performance. Aspects of participants' knowledge prior to using the DHTs included past experiences and associated emotions, user-system fit, health and disease knowledge, social influences and recommendations, and preexisting familiarity or trust toward the tool. Most participants therefore evaluated the performance of mHealth apps based on one or multiple outcomes: (1) the perceived severity and/or frequency of IBS symptoms before and after use, (2) changes to overall mood before and after use, (3) feelings toward the DHT, and (4) the ability of addressing concerns specific to the user. Online health resources, many of which served to disseminate disease knowledge and information, were evaluated differently by participants compared with apps. Participants instead focused on the accuracy, readability, accessibility, and perceived value of the information when assessing performance. One participant also noted the expectation-reality discrepancy among peers, contrasting high and unrealistic expectations of a fast, curative "fix" with DHTs designed for regular, long-term disease management.

Effort Expectancy

Most participants emphasized the importance of user-friendly DHT designs in successfully maintaining user engagement. Some described their experiences with poorly designed mHealth apps as frustrating or overwhelming (eg, rigid features, complicated functions requiring more user effort), and were subsequently discouraged from using the app further or to its fullest potential. Poor user design included features such as health monitoring functions requiring frequent data inputs or stringent timing, lacking customizable functions for personalization, and features with technical difficulties or lacking updates to address bugs. Conversely, apps with a more simplistic and seamless design reducing user effort, along with more customization capabilities, were viewed favorably. One participant considered it a priority to improve patient-centered aspects in DHT designs.

Participants unanimously agreed that patient engagement is critical for DHTs to ensure they are relevant and well-designed to address the needs of patients with IBS. Some participants considered that patients could engage in varying stages of development, such as the design and testing phases. The unanticipated burden of incorporating DHTs into their lifestyles—despite their intended use to improve patients' quality of life—further underscores the need for patient partners in the health research and technology space.

Social Influence

Many participants acknowledged their social circles as valuable support for their IBS diagnosis and management. They specified the type of support they received from their family and friends, specifically in navigating online resources, accessing the appropriate health care professionals, and recommending mHealth apps. However, some felt that certain DHT recommendations from their social circles were unhelpful or irrelevant to their circumstances because of limited knowledge or understanding of IBS. Others also were hesitant to fully embrace their support altogether, citing the social stigma around the condition.

In contrast, peer support groups were viewed favorably as a knowledgeable resource because of shared lived experiences with IBS. Most participants sought IBS support—often through social media platforms—for information regarding IBS-specific professional support, symptomatic management strategies, tools, and resources. One participant explained that peer recommendations could also help patients with IBS in decision-making about using mHealth apps by considering aspects or barriers from another perspective, such as evaluating the price relative to its value or benefits.

Facilitating Conditions

Participants' level of digital health literacy and competency with a device (eg, computer or mobile device) were identified as facilitating conditions contributing to user experiences with DHTs. Participants recognized the importance of digital health literacy in navigating, processing, and critiquing online health information. Many participants were considered to have higher digital health literacy because of their educational or professional backgrounds. Participants also acknowledged the role of computer and mobile competencies required for accepting and using DHTs, along with the significant learning curve for digital competency among older populations.

Hedonic Motivations

Most participants mentioned that their decision to adopt and use DHTs stemmed from a desire and willingness to learn about IBS and emphasized the importance of new and updated information available on these tools. Some participants had thoroughly researched and exhausted common IBS-related recommendations; DHTs that regularly provided new content were highly desirable and therefore maintained user engagement. Conversely, most participants struggled to fully engage with DHTs that lacked new content or variety, as they no longer felt the need to continue using these DHTs. Some participants also noted their interest in other IBS topics areas, such as mental health and community engagement.

Price Value

Participants frequently evaluated the cost-benefit trade-offs of DHTs—the monetary cost relative to their perceived values and benefits—in their decision-making process. Most participants first opted for services or apps that were free to download and use, but some also considered DHTs with paid business models. However, DHTs following a subscription business model were perceived by participants as the least desirable, with the accrued cost to maintain full access outweighing the potential benefits.

Habit

Participants correlated their habit or frequency of using DHTs with the DHTs' perceived utility, with greater frequency and established positive habits to be associated with symptomatic improvements and user satisfaction. Conversely, DHTs that were seldom used were deemed less useful, thus discouraging participants from using them even further.

The habit of DHT use depended on the features available, specifically reminders, and the digital platform. Some participants noted that customizable reminder functions prompted more consistent engagement from their experiences and were considered especially useful for certain DHT characteristics (eg, symptom and diet tracking and exercise programs); however, one participant mentioned that they did not need to be reminded to use their DHTs. Participants agreed that, given their frequent use of mobile devices (eg, smartphones), mHealth apps are more advantageous in maintaining user engagement. In contrast, participants perceived online websites or programs differently; they were viewed as cumbersome to locate, input login credentials, and search for the content. The integration of lifestyle with digital devices served as an antecedent factor that influenced participants' level of DHT acceptance and engagement.

Trust

In the context of the UTAUT2 model, trust in a DHT influences users' intention or willingness to use it and subsequently, whether the DHT is used for IBS (Figure 1). When navigating online health resources, participants positively associated medical or academic credibility with trust. Participants often referred to reputable medical websites (eg, Mayo Clinic and WebMD) for health information to improve their knowledge and IBS self-management. Academic and research institutions were also viewed favorably, and the involvement of a credible institution in the development of DHTs heavily weighted in favor of participants' decision-making. Some participants also acknowledged digital health literacy as a contributing factor toward trustworthiness of DHTs, as the ability to critically assess the information and sources provided greater confidence in recognizing credible websites and apps and placing trust in them.

Risk

In the acceptance and use of DHTs, participants weighed the potential risk associated with using DHTs as part of the trade-off with perceived benefits. Conflicting information, transparency concerns, and fear of financial and data exploitation were commonly raised by participants when navigating or using DHTs. For mHealth apps, concerns were primarily related to the handling and security of personal and health information. Regarding online resources, participants recognized the prevalence of conflicting information on the internet and the potential risks posed to unsuspecting users seeking clinical guidance. Participants also recounted their experiences with suspected scams, specifically websites that offered paid products and services to allegedly treat or cure IBS. One participant correlated their familiarity with the DHT and the DHTs' risk, associating less familiar DHTs with lower trustworthiness and greater risk.

Discussion

Principal Findings

Overview

Our findings highlight the importance of certain characteristics from both DHTs and users with IBS in determining intentions and success in navigating, adopting, and using DHTs, with many attributes aligning with the UTAUT2 framework [26]. Participants expressed an overarching theme of uncertainty towards DHTs in their abilities to ameliorate symptoms, which is based on their levels of and balance between perceived trust and risk. Other relevant user-specific attributes include digital health literacy and digital competency (facilitating conditions), desire for new knowledge (hedonic motivation), habitual use of DHTs (habit), and recommendations from social circles (social influence). DHT-specific characteristics include considerations around the design and features that influence perceived effectiveness (performance expectancy), user experience (effort expectancy), and value (price value). These considerations holistically illustrate the experiences and perspectives of patients with IBS using DHTs and their underlying and multifaceted influences. Addressing barriers and leveraging facilitators is essential for developing effective DHTs, improving user experience and health outcomes, and promoting innovation within the digital health care space [37-39].

Building Trust and Addressing Uncertainties of Digital Health Technologies

Trust is a juxtaposing force to fear of uncertainty and a critical component of successful implementation, adoption, and use of DHTs. However, the unprecedented evolution of the digital health landscape poses several challenges, including limited regulations, guidelines, and a lack of robust evidence base for efficacy, safety, and impact [40,41]. Furthermore, other limitations (eg, technological, costs, and transparency concerns) add to the DHTs' "black box" nature, which could discourage users from accepting or engaging with DHTs when they are not being fully informed.

In a 2018 scoping review, trust in digital health systems was influenced by several elements, including quality, efficiency, self-efficacy, accessibility, reputation and users' recommendations, and fear of data exploitation [42]. Many of these enabling and inhibiting elements of trust align with the key constructs of the applied UTATU2 model, further highlighting the complexity and interconnectedness of patient trust in adopting and using DHTs. Participants' trust in DHTs was also compromised by their fear and the potential risk of harm through scams, malicious digital attacks, and exploitation. Similarly, Catapan et al [43] found commercial and data exploitation concerns to undermine the trustworthiness of digital health in a cohort of patients with chronic kidney disease.

Participants' hesitancy toward DHTs must also be understood within the broader digital misinformation landscape. Research has shown that IBS-related content on online platforms, such as YouTube and TikTok, is made accessible in the absence of a vetting process, and can be inaccurate, promotional, or lack scientific grounding [44,45]. Likewise, online dietary guidance for IBS often suffers from poor readability and inconsistent quality [46]. Patients with chronic, fluctuating symptoms such as IBS might be particularly vulnerable to online "cures," unregulated wellness products, or misleading health claims [47]. In this context, skepticism reflects an adaptive response to navigating oversaturated but often unreliable digital health information. Recognizing health misinformation as an external driver of uncertainty reinforces the need for credible and evidence-based DHTs that patients can trust.

Reducing the uncertainties of DHTs and improving patient trust require a large-scale, coordinated, multidisciplinary approach involving technology developers, patients, researchers, health care professionals and institutions, and government bodies. First, improving the technological systems and processes of DHTs for quality, data security, and transparency is needed, with patients and researchers involved to assess effectiveness, utility, value, and feasibility [38]. Second, improving health care system infrastructures to better respond and adapt to DHTs will further facilitate general trust in DHTs among patients and health care professionals [38]. Lastly, stronger regulatory practices and guidelines should be implemented to improve transparency by ensuring that commercial and publicly available DHTs meet efficacy, safety, and privacy standards. While governments have the capacity and means to implement laws and regulatory measures, professional health associations and bodies could also play a pivotal role in supporting health care professionals in navigating and using DHTs alongside their patients. For example, Torous and Roberts [48] argued for the American Psychiatric Association to provide digital health–related resources and guidelines to patients and clinicians. However, while advancing transparency and patient trust is crucial, it is also important to avoid overregulation, which could subsequentially limit or delay DHT innovation, adoption, and use [48].

Existing Gaps in Digital Health Literacy

Digital health literacy is a factor in participants' decision-making on adopting or using DHTs (facilitating condition) and a crucial determinant of health—especially in one's ability to navigate and critically assess health information online and leverage the benefits of effective DHTs [49]. Our sample was highly educated, which may have influenced their ability to evaluate the quality of online sources effectively without additional support. However, digital inequality is an ongoing issue; patients with lower levels of digital health literacy are more likely to struggle with navigating and using DHTs, considering that high levels of reading comprehension are expected [39,46,50]. Limited understanding of content, benefits, and risks in adopting health technologies hinders patients' ability to make informed decisions about DHTs, and by extensions, to fully trust them.

Improving digital health literacy requires a collaborative approach involving patients, health care professionals, and policymakers. Strategies include the availability of educational or digital health literacy programs [51] and adapting online information and resources to a broader audience [52]. Technology developers should also tailor their content and design based on the target users' characteristics, such as digital health literacy levels, and improve transparency around privacy and security to mitigate harm and build trust. By facilitating the conditions in which patients are equipped with the skills and confidence to effectively use DHTs, they could fully leverage these available resources.

The Financial Barrier to Digital Health Technologies

Unsurprisingly, affordability was raised as a common barrier for participants to access support, especially mHealth apps with paid models or in-app purchases. While paid apps may be perceived more positively than free apps [37], affordability, or the lack thereof, could impact patients' adherence to them [53]. In addition to the existing socioeconomic burdens of IBS [54], patients could be especially more conscious of the financial trade-off of adopting and using cost-associated DHTs. Participants’ preferences for one-time payment models over subscription models aligns with similar findings from a patient cohort with rheumatoid arthritis [55], suggesting that certain payment models for DHTs could also influence patients’ decision-making process. Although Xie et al [56] suggested technology developers consider subscription plans to improve affordability and users’ willingness to pay, this approach might instead deter patients with chronic diseases requiring sustained DHT use. More research is warranted to better understand the preference of patients with IBS for DHTs with certain business models to ensure accessibility and acceptance.

The Value of Patient Engagement in Digital Health Technology Development

Content and user-centered designs were important aspects of DHT acceptance and use [57]. As part of their hedonic motivations, participants valued new and varied content in DHTs and considered limited and outdated content as a major deterrent. Health information–seeking behaviors are prevalent in people with IBS [58], which could be attributed to a desire and hope for relief or a cure [47]. Furthermore, IBS is a complex disorder involving individualized brain–gut interactions and requires personalized medicine [9]. As a result, the accuracy, variety, and customizability of the content are considered invaluable for patients to explore self-management strategies from a multidisciplinary perspective [59].

User-centered design is an iterative and collaborative process that requires a foundational understanding of the target users and the context of their intended use [60]. The outcome of such design is heavily dependent on comprehensive research and observations, feedback, and recommendations from lay individuals and experts at all stages of development (eg, initial design, evaluation, and user feedback) [60,61]. Patient engagement in digital health research and development is, therefore, fundamental to constructing a successful user-centered design in DHTs for optimal efficacy and impact [61].

However, patient engagement remains limited in digital health and technology innovation [62], given existing challenges around digital health literacy, accessibility, and trust and transparency concerns [39]. Careful planning and considerations are also needed to leverage patients’ inputs effectively; patients are most often involved in usability testing, where DHTs are mostly finalized which consequently leaves little flexibility to implement feedback [62]. Although there has been encouraging efforts to further engage patients with IBS in digital health research and implementation science [63], more is needed to fully ascertain the implications of DHTs and develop robust evidence specific to this population group.

Strengths and Limitations

This study is the first to explore the experiences of patients with IBS in navigating the digital landscape for self-management using the UTAUT2 model. Our findings contribute to the limited existing literature by providing insight into how DHTs can support IBS symptom management while addressing concerns related to distrust and digital literacy. We also identified key barriers and facilitators that impact patient experience and outcomes and provide recommendations for improvement based on patient perspectives.

This study also has several limitations. The transferability of our findings may be limited due to the small sample size and the presence of self-selection and sampling bias. First, our recruitment method used purposive sampling, where individuals decided whether to participate in the study. Our sample was highly educated, demonstrated high levels of digital health literacy, and showed great interest and motivation in health research and therefore was not representative of the general population with IBS. Second, all participants identified as women, which may be due to a combination of potential gender-related differences in IBS prevalence [64], mHealth app use [65], or the gender disproportion of the recruitment pool from previous IBS studies. This, however, poses the challenge of evaluating gender-based differences in this study, especially when gender is a moderating variable in both the original and expanded UTAUT models [26,33]. Third, all but one participant lived in a single Canadian province; their access or experience with health services and other non-DHT support may be different from other regions of Canada or worldwide. Lastly, coding saturation could not be fully attained in this study due to the insufficient sample size. While we have identified the most prevalent themes, having additional focus groups would further capture the themes comprehensively [30]. This would also avoid data bias and ensure the content validity, replicability, and transferability of the study [66]. Nevertheless, this study achieved its aim of building a deeper understanding and appreciation of patient experiences, perspectives, and decision-making around DHTs used for IBS that could serve as an antecedent for future studies and a guide in developing more patient-centered resources and supports.

Future Directions

Given the central role of digital literacy in our findings, future research should prioritize developing and evaluating educational or digital literacy programs specifically designed for patients with IBS. Identifying the unique barriers these patients face can guide the development of effective educational interventions. Additionally, longitudinal studies examining how enhanced digital literacy impacts the adoption and sustained use of digital tools would be valuable.

An emerging aspect in digital health is the rapidly expanding role of artificial intelligence and large language models (LLMs), such as ChatGPT, in supporting patient decision-making. Early evaluations suggest advanced LLMs can provide clinically aligned guidance, summarize complex medical concepts, and support initial health information–seeking more effectively than traditional search engines [67-69]. For patients with IBS, LLMs could serve as navigational tools that help filter misinformation, translate evidence-based information into patient-friendly language, and direct users toward credible DHTs or reputable resources. At the same time, concerns regarding bias, hallucinations, privacy, dependence on the quality and reliability of available training data and prompts, and lack of transparent regulatory oversight highlight that these tools remain imperfect and require careful integration into IBS self-management

[70-72]. Because LLM outputs are only as trustworthy as the data and sources on which they are trained, they may inadvertently reproduce inaccuracies or amplify misinformation. As AI-driven tools become more embedded in digital health, understanding how patients with IBS engage with them and how these tools can be responsibly implemented will become a critical future next step in this research. As AI-enabled tools and LLMs become increasingly integrated into consumer-health information-seeking behaviors, future work should also examine how these systems can support digital health literacy and help patients with IBS navigate an often-confusing online environment. Understanding how patients with IBS engage with these emerging technologies will be essential for designing safe and supportive digital self-management tools.

The evidence base on the feasibility and implementation of digital health technology for IBS is still lacking but necessary for real-world application. Understanding certain DHT features and attributes in patients' decision-making process could serve as a guide for future research and projects to understand and prioritize key objectives. Comprehensive economic evaluations should consider the direct and indirect costs of adopting DHTs in patients' self-management practices, as cost was considered a major barrier among participants.

Another fundamental step in this field is to improve patient engagement in health research and technology development. User-centered design was highlighted as a crucial component for the success of DHTs, and the need for greater involvement of patients in developing DHTs. Patient partners in future research and commercial development of DHTs, especially in the design and testing phases, will help ensure these tools are intuitive and meet patients' needs. Patient engagement will also be valuable in providing opportunities to quantify the significance of the identified influences in this study; this information could offer further insight into patients' decision-making process and how they weigh their considerations holistically.

Conclusion

This study highlights the complexities of the digital landscape in the context of IBS. While there is optimism toward DHTs in enhancing self-management, concerns about digital health literacy, trustworthiness, access, and cost must be addressed. With user-friendly, accessible, trusted, and cost-effective tools, the effectiveness and adoption of digital health solutions can be enhanced, and the digital health user experience optimized. It is evident that integrating DHTs in IBS self-management is a multifaceted, complex task that requires comprehensive and interdisciplinary collaboration among technology developers, health services, health professionals, and patients. More efforts are needed to address the identified barriers and leverage facilitators to improve health outcomes and quality of life for patients managing IBS or other chronic diseases.

Acknowledgments

We would like to express our gratitude to our patient research partner, Chantel Wicks, for her support and contributions to participant recruitment, and to student volunteer Darren Zhang for analyzing and summarizing the prefocus group surveys for this study.

Funding

This research was supported in full by the IMAGINE (Inflammation, Microbiome, and Alimentation: Gastro-Intestinal and Neuropsychiatric Effects), a Canadian Institutes for Health Research (CIHR) Strategy for Patient-Oriented Research (SPOR) Chronic Disease Network (funding reference number: 1715-000-001), with funding from several partners. No other financial arrangements were related to the research or assistance with manuscript preparation. DAM is supported by the Svare Chair in Health Economics, Value and Impact.

Data Availability

The datasets generated or analyzed during this study are not publicly available, as participants did not provide consent to share their individual data.

Authors' Contributions

AD and NH contributed equally as co–first authors. AD, DAM, MVM, JS, and NH conceptualized and designed the study. AD and NH collected the data. AD and NH analyzed the data. AD and NH wrote the first draft of the manuscript. All other authors provided revisions to subsequent drafts. All authors contributed to and approved the final manuscript.

Conflicts of Interest

DAM reports personal fees from the Office of Health Economics, Analytica and Novartis, and meeting expenses from ISPOR—The Professional Society for Health Economics and Outcomes Research—and Illumina. AD, MVM, JS, and NH disclose no conflicts of interest.

Multimedia Appendix 1

COREQ checklist.
[PDF File (Adobe PDF File), 481 KB - humanfactors_v13i1e75012_app1.pdf]

Multimedia Appendix 2
Pre-focus group survey of participants’ characteristics, IBS severity, and digital health tools used for IBS self-management.
[DOC File , 33 KB - humanfactors_v13i1e75012_app2.doc]

Multimedia Appendix 3
Semistructured focus group interview guide.
[DOC File , 31 KB - humanfactors_v13i1e75012_app3.doc]

Multimedia Appendix 4
Participant quotes of observed influences in considering DHT adoption and use for IBS patients in relation to key constructs based on the Expanded Unified Theory of Acceptance and Use of Technology (UTAUT2) model.
[DOCX File , 30 KB - humanfactors_v13i1e75012_app4.docx]

References

1. Kamiński M, Łoniewski I, Marlicz W. "Dr. Google, I am in pain"-global internet searches associated with pain: a retrospective analysis of google trends data. Int J Environ Res Public Health 2020;17(3):954 [FREE Full text] [doi: 10.3390/ijerph17030954] [Medline: 32033087]
2. Moreno-Ligero M, Moral-Munoz JA, Salazar A, Failde I. mHealth intervention for improving pain, quality of life, and functional disability in patients with chronic pain: systematic review. JMIR Mhealth Uhealth 2023;11:e40844 [FREE Full text] [doi: 10.2196/40844] [Medline: 36729570]
3. Black CJ, Ford AC. Global burden of irritable bowel syndrome: trends, predictions and risk factors. Nat Rev Gastroenterol Hepatol 2020;17(8):473-486 [FREE Full text] [doi: 10.1038/s41575-020-0286-8] [Medline: 32296140]
4. Lacy BE, Mearin F, Chang L, Chey WD, Lembo AJ, Simren M. Bowel disorders. Gastroenterology 2016;150(6):1393-1407. [doi: 10.1053/j.gastro.2016.02.031] [Medline: 27144627]
5. Shiha MG, Aziz I. Review article: physical and psychological comorbidities associated with irritable bowel syndrome. Aliment Pharmacol Ther 2021;54 Suppl 1:S12-S23. [doi: 10.1111/apt.16589] [Medline: 34927759]
6. Cong X, Perry M, Bernier KM, Young EE, Starkweather A. Effects of self-management interventions in patients with irritable bowel syndrome: systematic review. West J Nurs Res 2018;40(11):1698-1720 [FREE Full text] [doi: 10.1177/0193945917727705] [Medline: 28854852]
7. Flik CE, van Rood YR, de Wit NJ. Systematic review: knowledge and educational needs of patients with irritable bowel syndrome. Eur J Gastroenterol Hepatol 2015;27(4):367-371. [doi: 10.1097/MEG.0000000000000252] [Medline: 25874506]
8. Moayyedi P, Andrews CN, MacQueen G, Korownyk C, Marsiglio M, Graff L. Canadian Association of Gastroenterology clinical practice guideline for the management of irritable bowel syndrome (IBS). J Can Assoc Gastroenterol 2019;2(1):6-29 [FREE Full text] [doi: 10.1093/jcag/gwy071] [Medline: 31294724]
9. Lacy BE, Pimentel M, Brenner DM, Chey WD, Keefer LA, Long MD. ACG clinical guideline: management of irritable bowel syndrome. Am J Gastroenterol 2021;116(1):17-44. [doi: 10.14309/ajg.0000000000001036] [Medline: 33315591]
10. Kinsinger SW. Cognitive-behavioral therapy for patients with irritable bowel syndrome: current insights. Psychol Res Behav Manag 2017;10:231-237 [FREE Full text] [doi: 10.2147/PRBM.S120817] [Medline: 28790872]
11. Biesiekierski JR, Manning LP, Murray HB, Vlaeyen JWS, Ljótsson B, Van Oudenhove L. Review article: exclude or expose? The paradox of conceptually opposite treatments for irritable bowel syndrome. Aliment Pharmacol Ther 2022;56(4):592-605. [doi: 10.1111/apt.17111] [Medline: 35775328]
12. Shah ED, Salwen-Deremer JK, Gibson PR, Muir JG, Eswaran S, Chey WD. Comparing costs and outcomes of treatments for irritable bowel syndrome with diarrhea: cost-benefit analysis. Clin Gastroenterol Hepatol 2022;20(1):136-144.e31. [doi: 10.1016/j.cgh.2020.09.043] [Medline: 33010413]
13. Ankersen DV, Weimers P, Burisch J. Whats “App-ening”: the help of new technologies in nutrition in digestive diseases. Curr Opin Clin Nutr Metab Care 2017;20(5):426-431. [doi: 10.1097/MCO.0000000000000399] [Medline: 28768297]
14. Williams M, Barclay Y, Harper L, Marchant C, Seamark L, Hickson M. Feasibility, acceptability and cost-efficiency of using webinars to deliver first-line patient education for people with irritable bowel syndrome as part of a dietetic-led gastroenterology service in primary care. J Hum Nutr Diet 2020;33(6):758-766. [doi: 10.1111/jhn.12799] [Medline: 32816367]
15. McCrone P, Everitt H, Landau S, Little P, Bishop FL, O'Reilly G. Cost effectiveness of therapist delivered cognitive behavioural therapy and web-based self-management in irritable bowel syndrome: the ACTIB randomised trial. BMC Gastroenterol 2021;21(1):276 [FREE Full text] [doi: 10.1186/s12876-021-01848-9] [Medline: 34229619]

16. Ankersen DV, Carlsen K, Marker D, Munkholm P, Burisch J. Using eHealth strategies in delivering dietary and other therapies in patients with irritable bowel syndrome and inflammatory bowel disease. J Gastroenterol Hepatol 2017;32 Suppl 1:27-31. [doi: 10.1111/jgh.13691] [Medline: 28244677]
17. O'Keeffe M, Jansen C, Martin L, Williams M, Seamark L, Staudacher HM. Long-term impact of the low-FODMAP diet on gastrointestinal symptoms, dietary intake, patient acceptability, and healthcare utilization in irritable bowel syndrome. Neurogastroenterol Motil 2018;30(1):e13154. [doi: 10.1111/nmo.13154] [Medline: 28707437]
18. Hunt M, Miguez S, Dukas B, Onwude O, White S. Efficacy of Zemedy, a mobile digital therapeutic for the self-management of irritable bowel syndrome: crossover randomized controlled trial. JMIR Mhealth Uhealth 2021;9(5):e26152 [FREE Full text] [doi: 10.2196/26152] [Medline: 33872182]
19. Brenner DM, Ladewski AM, Kinsinger SW. Development and current state of digital therapeutics for irritable bowel syndrome. Clin Gastroenterol Hepatol 2024;22(2):222-234 [FREE Full text] [doi: 10.1016/j.cgh.2023.09.013] [Medline: 37743035]
20. D'Silva A, Hua N, Modayil MV, Seidel J, Marshall DA. Digital health interventions are effective for irritable bowel syndrome self-management: a systematic review. Dig Dis Sci 2025;70(2):644-664. [doi: 10.1007/s10620-024-08672-7] [Medline: 39402208]
21. Karkar R, Schroeder J, Epstein DA, Pina LR, Scofield J, Fogarty J. TummyTrials: a feasibility study of using self-experimentation to detect individualized food triggers. Proc SIGCHI Conf Hum Factor Comput Syst 2017;2017:6850-6863 [FREE Full text] [doi: 10.1145/3025453.3025480] [Medline: 28516175]
22. Tonkin-Crine S, Bishop FL, Ellis M, Moss-Morris R, Everitt H. Exploring patients' views of a cognitive behavioral therapy-based website for the self-management of irritable bowel syndrome symptoms. J Med Internet Res 2013;15(9):e2672 [FREE Full text] [doi: 10.2196/jmir.2672] [Medline: 24001787]
23. D'Silva A, Islam Z, Marshall DA, Vallance JK, Nasser Y, Rajagopalan V. Experiences of irritable bowel syndrome patients in a virtual yoga program: qualitative findings from a clinical trial. Dig Dis Sci 2024;69(1):169-179. [doi: 10.1007/s10620-023-08125-7] [Medline: 37914888]
24. 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;15(8):e2635 [FREE Full text] [doi: 10.2196/jmir.2635] [Medline: 23912839]
25. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
26. 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. [doi: 10.2307/41410412]
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;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
28. Morgan DL. Group composition, group size, and number of groups. In: Basic and Advanced Focus Groups. Thousand Oaks: SAGE Publications, Inc; 2019:49-62.
29. Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. Qual Health Res 2019;29(10):1483-1496 [FREE Full text] [doi: 10.1177/1049732318821692] [Medline: 30628545]
30. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. Field Methods 2016;29(1):3-22. [doi: 10.1177/1525822X16639015]
31. Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. Aliment Pharmacol Ther 1997;11(2):395-402. [doi: 10.1046/j.1365-2036.1997.142318000.x] [Medline: 9146781]
32. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
33. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
34. Zhao J, Li B, Sun J, Zeng X, Zheng J. Determinants of chronic disease patients' intention to use Internet diagnosis and treatment services: based on the UTAUT2 model. Front Digit Health 2025;7:1543428 [FREE Full text] [doi: 10.3389/fdgth.2025.1543428] [Medline: 40822908]
35. 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;10(1):e27095 [FREE Full text] [doi: 10.2196/27095] [Medline: 35040801]
36. Arfi WB, Nasr IB, Kondrateva G, Hikkerova L. The role of trust in intention to use the IoT in eHealth: application of the modified UTAUT in a consumer context. Technol Forecast Soc Change 2021;167:120688. [doi: 10.1016/j.techfore.2021.120688]
37. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. JMIR Mhealth Uhealth 2015;3(2):e4283 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]

38. Mumtaz H, Riaz MH, Wajid H, Saqib M, Zeeshan MH, Khan SE, et al. Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review. Front Digit Health 2023;5:1203945 [FREE Full text] [doi: 10.3389/fdgth.2023.1203945] [Medline: 37840685]
39. Livieri G, Mangina E, Protopapadakis ED, Panayiotou AG. The gaps and challenges in digital health technology use as perceived by patients: a scoping review and narrative meta-synthesis. Front Digit Health 2025;7:1474956 [FREE Full text] [doi: 10.3389/fdgth.2025.1474956] [Medline: 40212901]
40. Wells C, Spry C. An Overview of Smartphone Apps. Can J Health Technol 2022;2(2). [doi: 10.51731/cjht.2022.269]
41. Sittig DF, Wright A, Coiera E, Magrabi F, Ratwani R, Bates DW. Current challenges in health information technology-related patient safety. Health Informatics J 2020;26(1):181-189 [FREE Full text] [doi: 10.1177/1460458218814893] [Medline: 30537881]
42. Adjekum A, Blasimme A, Vayena E. Elements of trust in digital health systems: scoping review. J Med Internet Res 2018;20(12):e11254 [FREE Full text] [doi: 10.2196/11254] [Medline: 30545807]
43. Catapan SDC, Taylor ML, Scuffham P, Smith AC, Kelly JT. Improving consumer trust in digital health: a mixed methods study involving people living with chronic kidney disease. Digit Health 2025;11:20552076241312440 [FREE Full text] [doi: 10.1177/20552076241312440] [Medline: 39801583]
44. Jafri FI, Ancha N, Roy A, Ahmad S, Riaz H, Hyman C. TikTok and irritable bowel syndrome (IBS): a cross-sectional study of social media content quality. J Public Health (Berl.) 2023;33(3):653-658. [doi: 10.1007/s10389-023-02062-5]
45. Başmısırlı E, Kip M, Altun H, İnanç N. Evaluating the quality and reliability of YouTube videos providing nutritional recommendations for irritable bowel syndrome. J Hum Nutr Diet 2025;38(4):e70088. [doi: 10.1111/jhn.70088] [Medline: 40611462]
46. Alfaro-Cruz L, Kaul I, Zhang Y, Shulman RJ, Chumpitazi BP. Assessment of quality and readability of internet dietary information on irritable bowel syndrome. Clin Gastroenterol Hepatol 2019;17(3):566-567 [FREE Full text] [doi: 10.1016/j.cgh.2018.05.018] [Medline: 29800724]
47. Harvey JM, Sibelli A, Chalder T, Everitt H, Moss-Morris R, Bishop FL. Desperately seeking a cure: treatment seeking and appraisal in irritable bowel syndrome. Br J Health Psychol 2018;23(3):561-579 [FREE Full text] [doi: 10.1111/bjhp.12304] [Medline: 29508539]
48. Torous J, Roberts LW. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. JAMA Psychiatry 2017;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
49. van Kessel R, Wong BLH, Clemens T, Brand H. Digital health literacy as a super determinant of health: more than simply the sum of its parts. Internet Interv 2022;27:100500 [FREE Full text] [doi: 10.1016/j.invent.2022.100500] [Medline: 35242586]
50. Butcher CJ, Hussain W. Digital healthcare: the future. Future Healthc J 2022;9(2):113-117 [FREE Full text] [doi: 10.7861/fhj.2022-0046] [Medline: 35928188]
51. Busse TS, Nitsche J, Kernebeck S, Jux C, Weitz J, Ehlers JP, et al. Approaches to improvement of digital health literacy (eHL) in the context of person-centered care. Int J Environ Res Public Health 2022;19(14):8309 [FREE Full text] [doi: 10.3390/ijerph19148309] [Medline: 35886158]
52. Bywall KS, Norgren T, Avagnina B, Gonzalez MP, Andersson SW, IDEAHL consortium. Calling for allied efforts to strengthen digital health literacy in Sweden: perspectives of policy makers. BMC Public Health 2024;24(1):2666 [FREE Full text] [doi: 10.1186/s12889-024-20174-9] [Medline: 39350078]
53. Jakob R, Harperink S, Rudolf AM, Fleisch E, Haug S, Mair JL. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. J Med Internet Res 2022;24(5):e35371 [FREE Full text] [doi: 10.2196/35371] [Medline: 35612886]
54. Bosman MHMA, Weerts ZZRM, Snijkers JTW, Vork L, Mujagic Z, Masclee AAM. The socioeconomic impact of irritable bowel syndrome: an analysis of direct and indirect health care costs. Clin Gastroenterol Hepatol 2023;21(10):2660-2669 [FREE Full text] [doi: 10.1016/j.cgh.2023.01.017] [Medline: 36731587]
55. Chahal S, Biln N, Clarke B. Patient perspectives on a digital mobile health application for RA. Open Access Rheumatol 2021;13:63-71 [FREE Full text] [doi: 10.2147/OARRR.S296541] [Medline: 33935522]
56. Xie Z, Liu H, Or C. A discrete choice experiment to examine the factors influencing consumers' willingness to purchase health apps. Mhealth 2023;9:21. [doi: 10.21037/mhealth-22-39] [Medline: 37492118]
57. Molina-Recio G, Molina-Luque R, Jiménez-García AM, Ventura-Puertos PE, Hernández-Reyes A, Romero-Saldaña M. Proposal for the user-centered design approach for health apps based on successful experiences: integrative review. JMIR Mhealth Uhealth 2020;8(4):e14376 [FREE Full text] [doi: 10.2196/14376] [Medline: 32319965]
58. Shin A, Ballou S, Camilleri M, Xu H, Lembo A. Information- and health-care seeking behaviors in patients with irritable bowel syndrome. Clin Gastroenterol Hepatol 2020;18(12):2840-2842 [FREE Full text] [doi: 10.1016/j.cgh.2019.09.020] [Medline: 31546055]
59. Staudacher HM, Black CJ, Teasdale SB, Mikocka-Walus A, Keefer L. Irritable bowel syndrome and mental health comorbidity - approach to multidisciplinary management. Nat Rev Gastroenterol Hepatol 2023;20(9):582-596 [FREE Full text] [doi: 10.1038/s41575-023-00794-z] [Medline: 37268741]

60. Wiljén A, Chaplin JE, Crine V, Jobe W, Johnson E, Karlsson K. The development of an mHealth tool for children with long-term illness to enable person-centered communication: user-centered design approach. JMIR Pediatr Parent 2022;5(1):e30364 [FREE Full text] [doi: 10.2196/30364] [Medline: 35258466]
61. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015;22(6):754-756 [FREE Full text] [doi: 10.1111/acem.12692] [Medline: 25997375]
62. Baines R, Bradwell H, Edwards K, Stevens S, Prime S, Tredinnick-Rowe J. Meaningful patient and public involvement in digital health innovation, implementation and evaluation: a systematic review. Health Expect 2022;25(4):1232-1245 [FREE Full text] [doi: 10.1111/hex.13506] [Medline: 35526274]
63. Olson JL, Palumbo A, Desjardins C, Wicks C, Bhopa S, Cheyne K. Perspectives on the sustained engagement with digital health tools: protocol for a qualitative interview study among people living with inflammatory bowel disease or irritable bowel syndrome. BMJ Open 2024;14(11):e089220 [FREE Full text] [doi: 10.1136/bmjopen-2024-089220] [Medline: 39521469]
64. Kim YS, Kim N. Sex-gender differences in irritable bowel syndrome. J Neurogastroenterol Motil 2018;24(4):544-558 [FREE Full text] [doi: 10.5056/jnm18082] [Medline: 30347934]
65. 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]
66. Fusch P, Ness L. Are we there yet? Data saturation in qualitative research. Qual Rep 2015;20(9). [doi: 10.46743/2160-3715/2015.2281]
67. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW. Large language models encode clinical knowledge. Nature 2023;620(7972):172-180 [FREE Full text] [doi: 10.1038/s41586-023-06291-2] [Medline: 37438534]
68. Bopche R, Gustad LT, Afset JE, Ehrnström B, Damås JK, Nytrø Ø. Leveraging explainable artificial intelligence for early prediction of bloodstream infections using historical electronic health records. PLOS Digit Health 2024;3(11):e0000506. [doi: 10.1371/journal.pdig.0000506] [Medline: 39541276]
69. Chen J, Ge X, Yuan C, Chen Y, Li X, Zhang X. Comparing orthodontic pre-treatment information provided by large language models. BMC Oral Health 2025;25(1):838 [FREE Full text] [doi: 10.1186/s12903-025-06246-1] [Medline: 40437500]
70. Cirillo D, Catuara-Solarz S, Morey C, Guney E, Subirats L, Mellino S. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. NPJ Digit Med 2020;3(1):81 [FREE Full text] [doi: 10.1038/s41746-020-0288-5] [Medline: 32529043]
71. Yang S, Jing M, Wang S, Huang Z, Wang J, Kou J. Building trustworthy large language model-driven generative recommender system for healthcare decision support: A scoping review of corpus sources, customization techniques, and evaluation frameworks. Artif Intell Med 2026;171:103310. [doi: 10.1016/j.artmed.2025.103310] [Medline: 41313967]
72. Geroimenko V. Generative AI hallucinations in healthcare: a challenge for prompt engineering and creativity. In: Geroimenko V, editor. Human-Computer Creativity: Generative AI in Education, Art, and Healthcare. Switzerland: Springer Nature; 2025:321-335. [doi: 10.1007/978-3-031-86551-0_17]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
DHT: digital health technology
FODMAP: fermentable oligosaccharides, disaccharides, monosaccharides, and polyols
IBS: irritable bowel syndrome
IBS-SSS: IBS Symptom Severity Score
LLM: large language model
mHealth: mobile health
UTAUT2: Expanded Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 26.Mar.2025; peer-reviewed by J Soldera; comments to author 16.Oct.2025; accepted 02.Jan.2026; published 02.Feb.2026.

Please cite as:
D'Silva A, Hua N, Modayil MV, Seidel J, Marshall DA
Navigation, Adoption, and Use of Digital Health Technologies for Irritable Bowel Syndrome Self-Management: Focus Group Study of Patient Experience and Decision-Making
JMIR Hum Factors 2026;13:e75012
URL: https://humanfactors.jmir.org/2026/1/e75012
doi:10.2196/75012
PMID:

Correction: Auxiliary Teaching and Student Evaluation Methods Based on Facial Expression Recognition in Medical Education

Xueling Zhu[1,2], MSE; Roben A Juanatas[1], PhD

[1]College of Computing and Information Technologies, National University, 551 Mariano Fortunato Jhocson Street, Sampaloc, Manila, Philippines
[2]College of Big Data and Artificial Intelligence, Anhui Xinhua University, Hefei, Anhui, China

Corresponding Author:
Xueling Zhu, MSE
College of Computing and Information Technologies, National University, 551 Mariano Fortunato Jhocson Street, Sampaloc, Manila, Philippines

Related Article:

Correction of: https://humanfactors.jmir.org/2025/1/e72838

Abstract

(JMIR Hum Factors 2026;13:e89669) doi:10.2196/89669

In “Auxiliary Teaching and Student Evaluation Methods Based on Facial Expression Recognition in Medical Education” [1], the authors noted one error in the author affiliation list.

A second affiliation has now been added for author XZ, as follows:

> *College of Big Data and Artificial Intelligence, Anhui Xinhua University, Hefei, Anhui, China*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Zhu X, Juanatas RA. Auxiliary teaching and student evaluation methods based on facial expression recognition in medical education. JMIR Hum Factors 2025 May 22;12:e72838. [doi: 10.2196/72838] [Medline: 40402552]

Submitted 16.Dec.2025; this is a non–peer-reviewed article;accepted 16.Dec.2025; published 14.Jan.2026.

Please cite as:
Zhu X, Juanatas RA
Correction: Auxiliary Teaching and Student Evaluation Methods Based on Facial Expression Recognition in Medical Education
JMIR Hum Factors 2026;13:e89669
URL: https://humanfactors.jmir.org/2026/1/e89669
doi:10.2196/89669

Acceptability of a Digital Care App in Patients Undergoing Hip and Knee Arthroplasty: Prospective Cohort Study

Yacine Louni[1], MDCM; Matthew Laroche[1], BCom; Abdulrhman Alnasser[2], MD; Mohammad Abuhaneya[2], MD; Eric Belzile[3], MSc; Sandhya Baskaran[3], MSc; Jennifer Mutch[2], MDCM, FRCSC; Anthony Albers[2], MDCM, FRCSC

[1]Faculty of Medicine, McGill University, Montreal, QC, Canada
[2]Division of Orthopaedic Surgery, McGill University, St Mary’s Hospital, 3830 Avenue Lacombe, Montreal, QC, Canada
[3]St Mary's Hospital Research Center, St Mary's Hospital, Montreal, QC, Canada

Corresponding Author:
Anthony Albers, MDCM, FRCSC
Division of Orthopaedic Surgery, McGill University, St Mary’s Hospital, 3830 Avenue Lacombe, Montreal, QC, Canada

Abstract

Background: Mobile health (mHealth) apps have become more commonly used in orthopedics. However, for these apps to be efficient, patients should be willing to use them, making it essential to understand patients’ perspectives of mHealth interventions.

Objective: The aim of this single-center, intent-to-treat, preoperative single-cohort study of 100 patients was to evaluate the acceptability of mymobility (Zimmer-Biomet), an mHealth app designed for the postoperative care of total hip arthroplasty (THA) and total knee arthroplasty (TKA).

Methods: We measured acceptability using the theoretical framework for acceptability (TFA) preoperatively and at 3 months post operation. We also measured satisfaction with app use postoperatively using the Usefulness, Satisfaction, and Ease of Use questionnaire as well as patient-reported outcome measures preoperatively and postoperatively using the Oxford hip and knee scores and the visual analog scale for pain. Patients included were 18 years or older; underwent unilateral primary total hip, total knee, or partial knee arthroplasty; spoke and read French or English; and had a smartphone with internet access. Participants used mymobility in addition to standard government-funded physiotherapy.

Results: The preoperative overall TFA result was 4.2 out of 5, but results decreased significantly postoperatively. There was higher self-efficacy in preoperative TFAs with higher education, and lower acceptability in postoperative TFAs with TKA. The Usefulness, Satisfaction, and Ease of Use questionnaire revealed a good level of satisfaction with the use of the app. Patient-reported outcome measures showed earlier improvement in THA (31.2 d) than in TKA (89.4 d), whereas the visual analog scale showed a rapid decrease in pain with both procedures. Only 1 patient expressed privacy concerns with the use of the app.

Conclusions: There was a good level of acceptability with the use of mymobility for the postoperative management of THA and TKA, although acceptability decreased postoperatively. This decrease could signify high expectations toward the app preoperatively or higher than expected difficulty and pain in the early postoperative period. Acceptability tended to increase with higher education and decrease with TKA. These trends are consistent with prior literature and constitute a potential gap to address for app developers. The influence of the natural recovery process on acceptability remains unclear. Future studies could explore this gap by comparing results in cohorts using the app to cohorts with standard care.

(JMIR Hum Factors 2026;13:e79682) doi:10.2196/79682

KEYWORDS

mobile health; orthopedics; patient satisfaction; telemedicine; telerehabilitation; total joint arthroplasty

Introduction

In Canada, there were 60 705 hip and 70 379 knee replacements performed in the inpatient setting in 2023 - 2024 [1]. These procedures were the third and second most common inpatient surgeries, and these numbers are expected to at least double in the United States (and likely Canada) over the next 15 years due to an aging population [2,3]. Another growing sector in health care is the use of smartphones for patient care. An estimated 71% of the world population uses a smartphone in 2024, and there are more than 350 000 mobile health (mHealth) apps currently available [4,5]. This increasing use of mHealth apps is also present in various areas of orthopedics, where numerous studies have shown that they are efficient and convenient, that they promote patient adherence and engagement with their treatment, and that they facilitate rehabilitation while promoting positive outcomes [6-13].

Although mHealth has the potential to improve outcomes, patients’ attitudes toward these apps must be assessed, as they can only be effective if patients are willing to use them. User acceptability is defined as “a multifaceted construct that reflects the extent to which people delivering or receiving health care intervention consider it to be appropriate based on anticipated or experienced cognitive and emotional responses to the intervention” [14]. Acceptability is an important measure, as it not only provides information on patients’ experience with mHealth apps, but also gives an insight into their cognitive and emotional response [14]. This helps determine whether patients genuinely perceive them as appropriate and valuable.

Recent studies investigated patients’ perspectives of mHealth interventions in the context of postoperative hip and knee arthroplasty care. These studies found a high level of engagement and increased compliance with treatment exercises when using mHealth interventions [13,15-18]. Other findings include participants being satisfied with these interventions while also finding them easy to use and engaging, and providing better connection to their treating team [13,16,19,20]. These results are promising and could signify that patients are willing to use mHealth apps for their rehabilitation. However, these studies either did not assess acceptability or did so through surveys that were researcher-created or not primarily designed to measure acceptability. According to Perski et al [21], this reduces clarity regarding which aspects of acceptability are satisfied, and which are not. Additionally, high engagement does not necessarily reflect patients’ perceptions of the intervention. Engagement might be driven by external factors, such as study participation, financial incentives, or free access to the devices and apps, while patients might still find the app to be time-consuming, repetitive, and burdensome.

Measuring acceptability, especially preoperatively, is therefore relevant as it provides a better perspective of patients’ opinions prior to and beyond use, and because it “may usefully be considered an emergent property […] which in turn influences (and is influenced by) user engagement and intervention effectiveness” [21].

We started using the mymobility app (Zimmer Biomet, Warsaw, and Indiana) in 2022 in our community hospital to provide postoperative care to patients receiving outpatient total hip arthroplasty (THA) and total knee arthroplasty (TKA) [22]. This app, available on iOS and Android, currently stands at more than 50 000 downloads on the Google Play Store and is amongst the most downloaded when compared to other apps with the same objective [23]. In this context, our team sought to investigate the acceptability of mymobility, an mHealth app specifically designed for the postoperative care of THA and TKA, using a standardized framework.

Methods

Study Design

The primary aim of this single-center, intent-to-treat (ITT), single-cohort study was to evaluate pre- and postoperative patient acceptability with the use of the mymobility app for the postoperative care of THA and TKA. The secondary outcomes were postoperative patient satisfaction with the app and pre- and postoperative patient-reported outcome measures (PROMs).

Data Collection

Data collected for this study are presented in Multimedia Appendix 1. Demographic data were obtained from patient charts. PROMs were collected either directly through the mymobility app or by one of the authors contacting participants by phone or email. Details of the specific PROMs are provided later in this section. All data were stored securely and deidentified prior to analysis.

Ethical Considerations

This study received ethics approval from the Research Ethics Board of the Montreal West Island’s Integrated University Health and Social Services Centers (Biomedical Subcommittee), affiliated with St. Mary’s Hospital, Montreal (IRB# 2024-954). All participants provided informed consent prior to participation. Study data, including personal information, remained confidential and were deidentified. No secondary analyses were conducted. Study participants received no financial compensation, but access to the mymobility app was given for free. No identifiable images of participants were included in this manuscript. Under the Personal Information Protection and Electronic Documents ACT, the data collected by mymobility could not be used or shared with third parties, nor could cookies be collected, unless participants gave their explicit consent within the app [24,25]. If consent was given, data could be used or shared by the app following anonymization, with consent revocable at any time. Participants were not required to give consent to data use or sharing, or to cookie collection to use the app and be part of the study. This was made clear to participants as part of their informed consent.

App Description

Mymobility is an app that is accessible either through a smartphone or an Apple Watch. The app features educational content on pre- and postoperative care, exercise routines with metrics to track progress, AI-assisted gait speed analysis to determine the level of recovery, in-app PROMs, and telemedicine with the opportunity to send SMS text messages, videos, or pictures to health care providers (see Multimedia Appendix 2). The app engages patients in their recovery by sending notifications and reminders if users accept them while also keeping track of progress made. It also provides patients with basic and important smartphone functions to use (eg, Wi-Fi) for those with low digital literacy.

Questionnaires

Acceptability

Acceptability was measured using the theoretical framework for acceptability (TFA) questionnaire developed by Sekhon et al [26], which was adapted to our research question (see Multimedia Appendix 3). The questionnaire comprises 8 items answered by participants preoperatively and at 3 months post operation by phone or email. Each statement was evaluated using a 5-point Likert scale: scores 1 to 2 were classified as negative, 3 as neutral, and 4 - 5 as positive. The mean score was calculated for each of the 8 items, as well as an overall score.

Satisfaction

Patient satisfaction was measured using the Usefulness, Satisfaction, and Ease of Use (USE) questionnaire at 3 months post operation (see Multimedia Appendix 4) [27]. Satisfaction was defined as “the net feeling of pleasure or displeasure that results from aggregating all the benefits that a person hopes to receive from interaction with the information system” [27]. The questionnaire was answered by phone or email. Statements were evaluated using a 7-point Likert scale: scores 1 to 3 were classified as negative, 4 as neutral, and 5 to 7 as positive. The mean score was then calculated for each of the 4 sections of the questionnaire. Two additional questions assessing satisfaction were asked and scored on a 5-point Likert scale (see Multimedia Appendix 5).

Patient-Reported Outcome Measures

PROMs were measured using the Oxford Hip Score (OHS) and the Oxford Knee Score (OKS) as well as the visual analog scale (VAS) for pain (scored from 0 to 10) preoperatively and at 1, 3, and 6 months post operation (see Multimedia Appendices 6 and 7) [28-30]. OHS and OKS were answered directly within the app, whereas the VAS was answered by phone or email.

Recruitment

Patients were included in the study if they were booked for unilateral primary THA or TKA, were clinically suited for telehealth care (characterized as age ≥18 y, speak and read French or English, and have a smartphone or a tablet), had internet access, and provided informed consent to participate in the study. Patients were excluded if they did not meet the inclusion criteria, if they were unable or unwilling to provide informed consent, if they were undergoing emergency THA or TKA, or if they were undergoing revision surgeries. Participants had access to the app in addition to normal postoperative care, which included physiotherapy and postoperative follow-up visits with an orthopedic surgeon.

Patients were approached by a member of the research team during their preoperative assessment visit in the clinic. Eligible patients were provided with an informed consent form detailing the study. Once enrolled, a 1-year free access to the mymobility app was provided to the participants by the treating physician or a member of the research team, after which an email with instructions to download the app was sent. Participants were then required to download the app and agree to the terms and conditions to use it [24].

In the end, 100 participants were included in this study and had a profile created in the app. Participants who did not download the app preoperatively were sent a reminder email and given a few days following surgery, after which they were considered as nondownloaders. In the end, 75 participants downloaded the app and 25 did not. Only participants who downloaded the app were provided with the study questionnaires, as the goal was to assess acceptability, satisfaction, and PROMs with use of the app. Patients who did not download the app remained in the study in the context of the ITT format. However, since deciding not to download the app could signify low acceptability, we contacted participants who did not download the app by phone or email to ask the reason behind their decision.

Data Analysis

Study data was analyzed using the ITT principle. Data analysis was conducted and reported using STATA 17.0 and R software. Statistical significance was defined as a *P* value <.05.

Results

The recruitment process is summarized in Figure 1. Baseline characteristics of our population can be found in Table 1.

Figure 1. Study enrollment, participation, and questionnaire distribution flow diagram. PROM: patient-reported outcome measure; TFA: theoretical framework for acceptability; USE: Usefulness, Satisfaction, and Ease of Use.

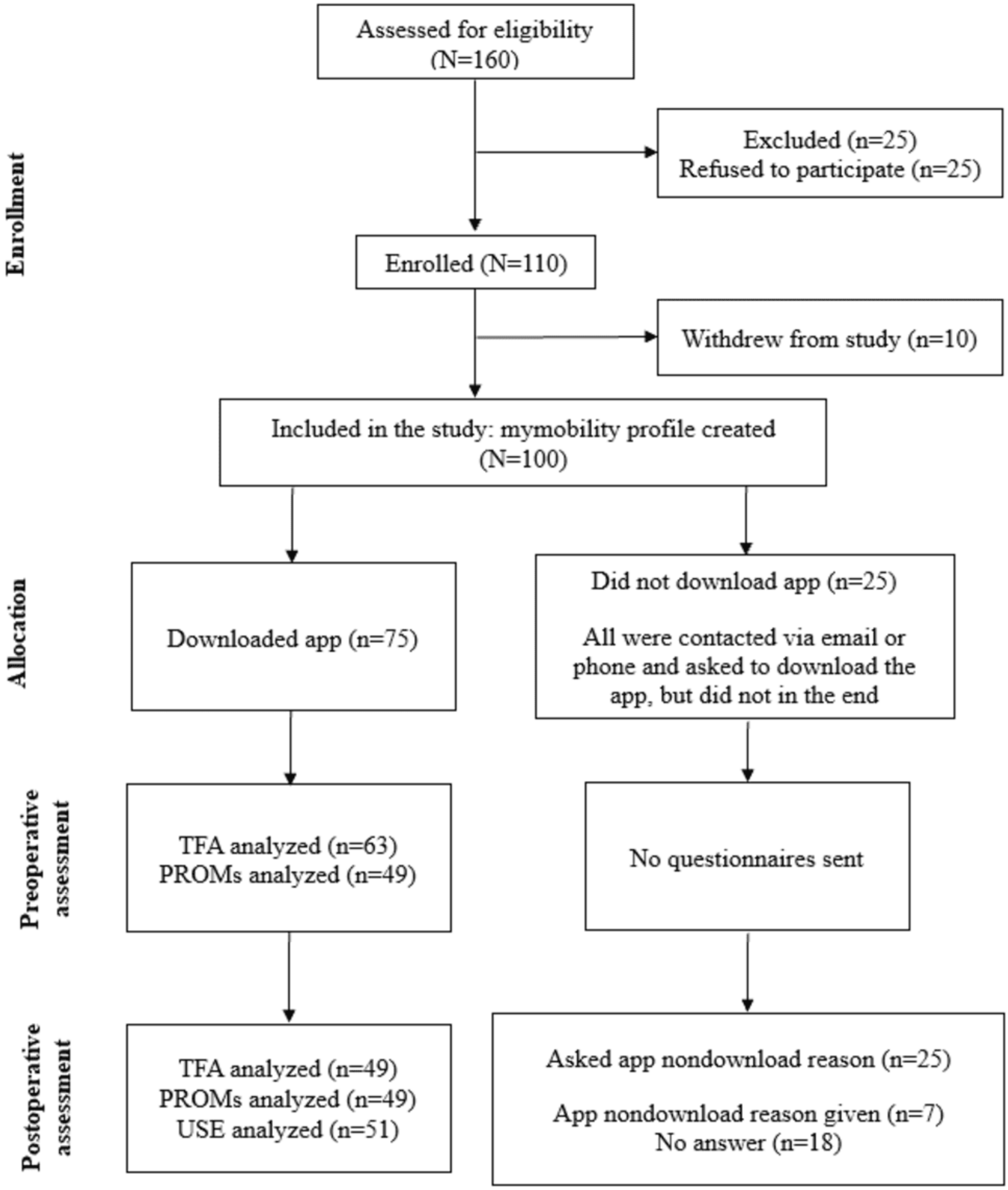


Table . Baseline demographic and clinical characteristics of the study population.

Baseline variables	Participants (N=100)
Female, n (%)	51 (52)
Missing	1 (1)
Age (years), mean (SD)	67.0 (10.7)
Ethnicity, n (%)	
Caucasian	75 (82)
Asian	9 (10)
Arab	3 (3)
Other	5 (5)
Missing	8 (8)
Education, n (%)	
Graduate	20 (21)
Undergraduate	27 (29)
Postsecondary (nonuniversity)	25 (27)
Primary/secondary	22 (23)
Missing	6 (6)
Employment, n (%)	
Full/part time	31 (34)
Retired	56 (61)
Unemployed	5 (5)
Missing	8 (8)
Language used, n (%)	
English	64 (69)
French	23 (25)
Both	6 (6)
Missing	7 (7)
Comorbidities, n (%)	
Yes[a]	39 (39)
None	61 (61)
Hypertension	18 (31)
Diabetes	10 (17)
Obesity	6 (10)
Thyroid	4 (7)
Cholesterol	4 (7)
Other	17 (29)
Procedure, n (%)	
TKA[b]	55 (55)
THA[c]	42 (42)
PKA[d]/UKA[e]	3 (3)
Income (CAD $[f]), n (%)	
<25,000	8 (9)
25,000 - 50,000	15 (17)

Baseline variables	Participants (N=100)
50,000 - 75,000	10 (11)
>100,000	24 (27)
Prefer not to say/unknown	13 (15)
Missing	11 (11)
Area of residency, n (%)	
Urban	72 (77)
Rural	22 (23)
Missing	6 (6)

[a] Participants who answered “Yes” can have multiple comorbidities.

[b] TKA: total knee arthroplasty.

[c] THA: total hip arthroplasty.

[d] PKA: partial knee arthroplasty.

[e] UKA: unicompartmental knee arthroplasty.

[f] A currency exchange rate of CAD $1=US $0.75 is applicable.

Acceptability

The overall score for the preoperative TFA questionnaires was 4.2 (SD 0.6; Table 2). We found a statistically significant increase in self-efficacy with a university level of education compared to non-university (*P*=.04), but no difference related to age, sex, or employment (Figure 2). There was a statistically significant decrease between pre- and postoperative overall TFA (*P*=.007), as well as in pre- and postoperative perceived effectiveness (*P*=.01) and self-efficacy (*P*=.008; Table 3). Furthermore, when looking at TFA results by procedure, there was a statistically significant decrease with TKA in multiple TFA items, including overall acceptability (*P*=.008; Figure 3). There were no statistically significant differences between pre- and postoperative overall TFA nor in any of the TFA items with THA.

Figure 2. Preoperative TFA results (all items) for the overall population and by education level. TFA: theoretical framework for acceptability.

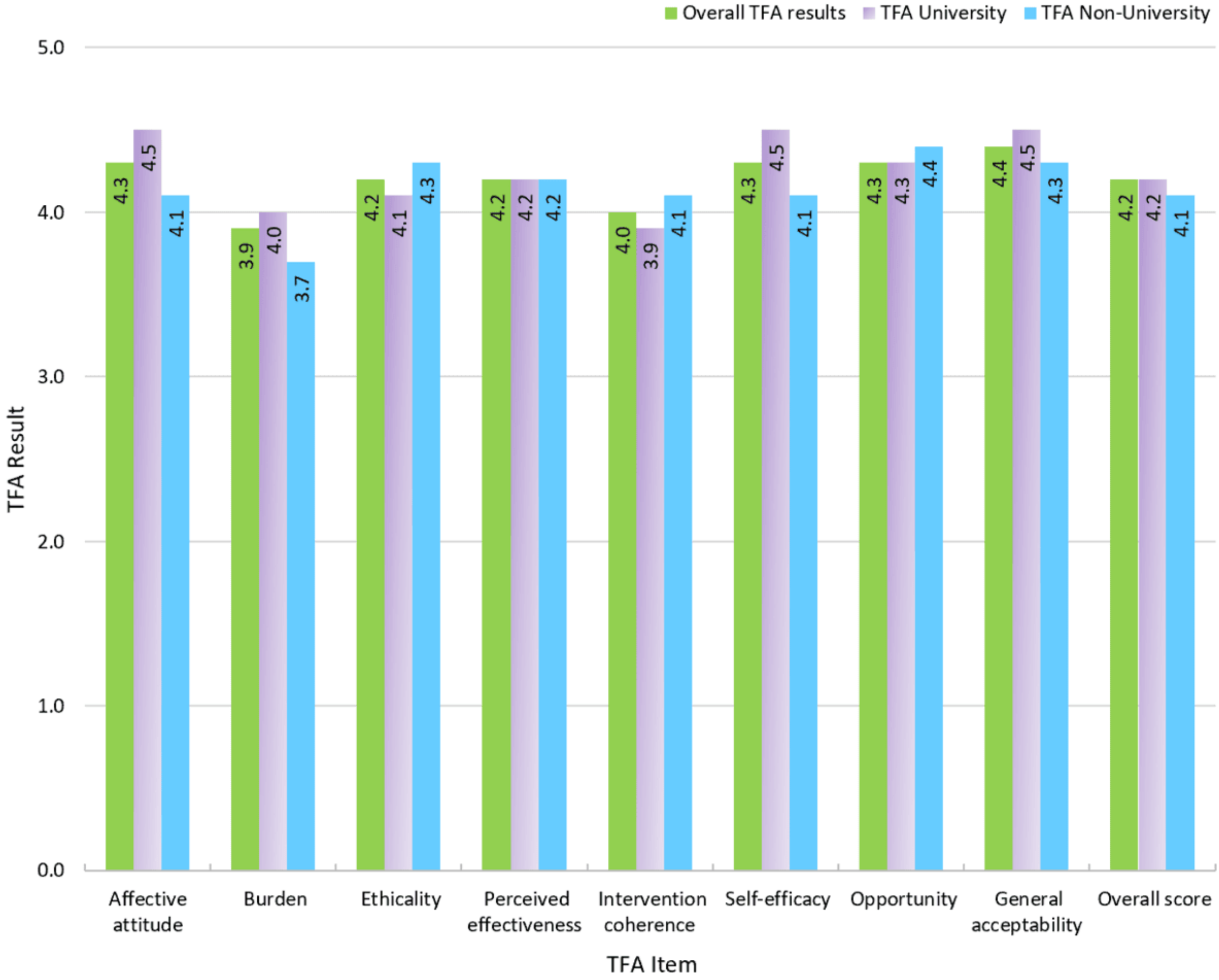


Figure 3. Comparison between pre- and postoperative overall TFA results for the overall population and by procedure. TFA: theoretical framework for acceptability; THR: total hip replacement; TKR: total knee replacement.

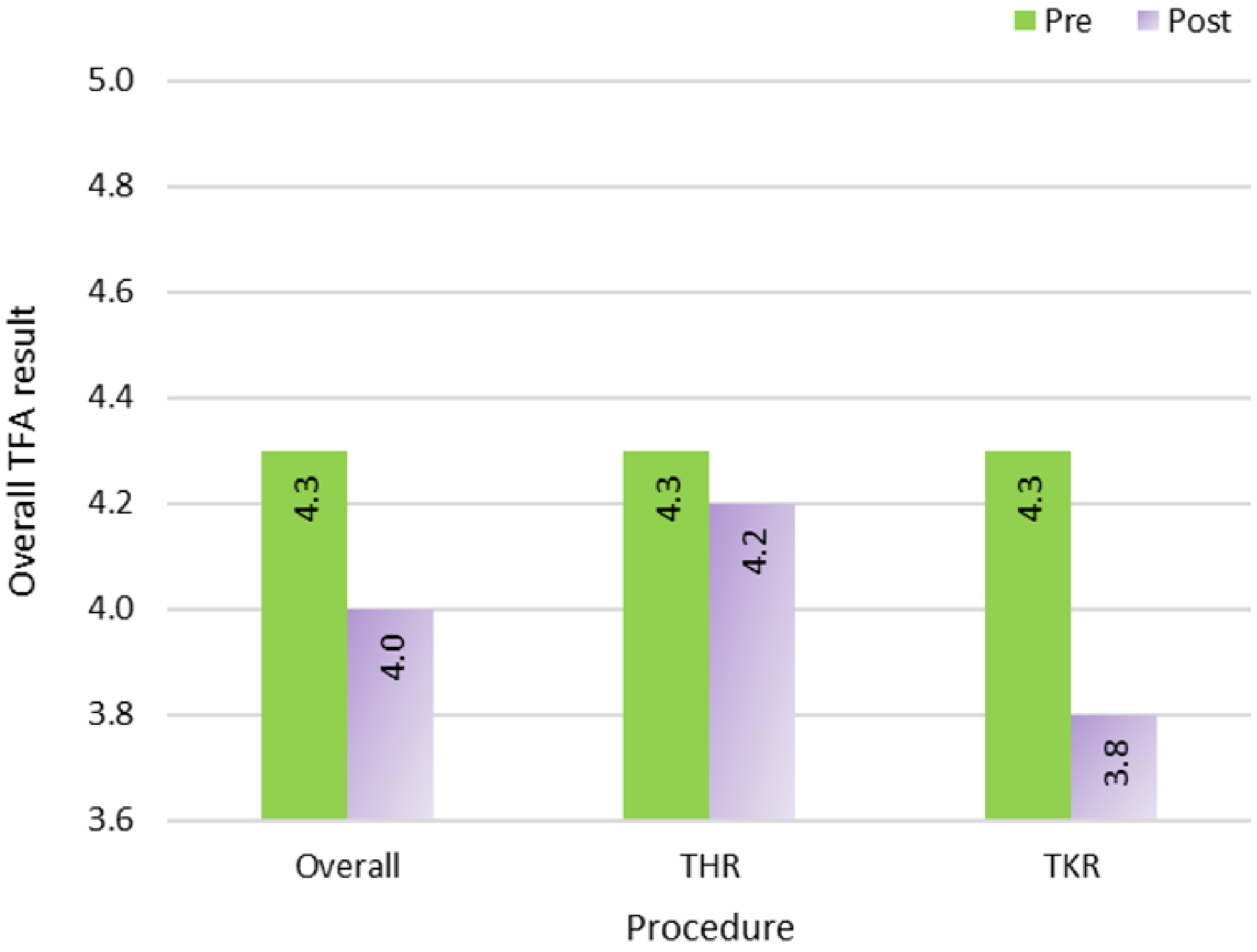


Table . Preoperative TFA[a] results and subgroup analysis by demographic category[b].

TFA items	Overall (n=63), mean (SD)	Age			Sex			Education			Employment		
		<70 (n=34), mean (SD)	≥70 (n=29), mean (SD)	*P* value (*t* test)	Male (n=33), mean (SD)	Female (n=30), mean (SD)	*P* value (*t* test)	Nonuniversity (n=28), mean (SD)	University (n=32), mean (SD)	*P* value (*t* test)	Not working (n=36), mean (SD)	Working (n=23), mean (SD)	*P* value (*t* test)
Affective attitude	4.3 (0.9)	4.3 (0.8)	4.4 (0.8)	.84	4.4 (0.8)	4.1 (1.0)	.21	4.1 (1.0)	4.5 (0.8)	.08	4.4 (0.8)	4.3 (1.1)	.67
Burden	3.9 (0.9)	3.9 (1.0)	3.8 (0.9)	.63	4.0 (0.7)	3.9 (1.1)	.47	3.7 (1.2)	4.0 (0.7)	.20	3.8 (1.0)	4.0 (0.9)	.63
Ethicality	4.2 (1.0)	4.1 (1.1)	4.3 (1.1)	.32	4.1 (1.1)	4.3 (0.8)	.34	4.3 (0.9)	4.1 (1.1)	.54	4.1 (1.1)	4.4 (0.8)	.22
Perceived effectiveness	4.2 (0.8)	4.1 (1.0)	4.3 (0.6)	.23	4.2 (0.9)	4.2 (0.9)	.96	4.2 (0.8)	4.2 (0.9)	.98	4.2 (0.7)	4.3 (1.2)	.55
Intervention coherence	4.0 (0.9)	3.8 (1.1)	4.3 (0.6)	.05	4.0 (0.9)	4.0 (1.0)	.99	4.1 (0.7)	3.9 (1.1)	.49	4.1 (0.8)	3.8 (1.2)	.21
Self-efficacy	4.3 (0.7)	4.3 (0.9)	4.3 (0.5)	.93	4.3 (0.8)	4.3 (0.7)	.99	*4.1 (0.9)*	*4.5 (0.6)*	*.04*	4.3 (0.5)	4.3 (1.1)	.89
Opportunity	4.3 (0.9)	4.2 (1.0)	4.5 (0.7)	.16	4.3 (0.9)	4.3 (0.8)	.88	4.4 (0.7)	4.3 (0.9)	.40	4.4 (0.7)	4.2 (1.0)	.27
General acceptability	4.4 (0.5)	4.3 (0.6)	4.4 (0.5)	.74	4.3 (0.5)	4.4 (0.6)	.87	4.3 (0.5)	4.5 (0.5)	.09	4.3 (0.5)	4.5 (0.6)	.40
Overall score[c]	4.2 (0.6)	4.1 (0.6)	4.3 (0.5)	.22	4.2 (0.5)	4.2 (0.7)	.78	4.1 (0.6)	4.2 (0.5)	.55	4.2 (0.5)	4.2 (0.6)	.94

[a] TFA: theoretical framework for acceptability.

[b] Statistically significant results are given in italic.

[c] Burden, ethicality, and opportunity costs were reversed to compute the overall TFA score.

Table . Change in TFA[a] from preoperative to 3 months post operation[b].

TFA items	Overall (n=49)			TKA[c] (n=27)			THA[d] (n=22)		
	Pre, mean (SD)	Post, mean (SD)	*P* value (*t* test)	Pre, mean (SD)	Post, mean (SD)	*P* value (*t* test)	Pre, mean (SD)	Post, mean (SD)	*P* value (*t* test)
Affective attitude	4.5 (0.6)	4.2 (1.3)	.09	4.5 (0.6)	4.1 (1.3)	.22	4.6 (0.7)	4.3 (1.3)	.26
Burden	4.0 (0.8)	4.0 (1.1)	.91	3.9 (1.0)	3.8 (1.3)	.70	4.1 (0.7)	4.3 (0.8)	.36
Ethicality	4.3 (0.9)	3.9 (1.3)	.11	*4.4 (0.9)*	*3.6 (1.4)*	*.04*	4.2 (0.9)	4.3 (1.1)	.89
Perceived effectiveness	*4.3 (0.8)*	*3.8 (1.3)*	*.01*	*4.3 (0.7)*	*3.7 (1.4)*	*.02*	4.2 (0.9)	4.0 (1.2)	.28
Intervention coherence	4.1 (0.9)	3.8 (1.2)	.09	*4.2 (0.8)*	*3.6 (1.3)*	*.01*	3.9 (1.1)	4.0 (1.0)	.69
Self-efficacy	*4.4 (0.5)*	*4.0 (1.2)*	*.008*	*4.4 (0.5)*	*3.8 (1.4)*	*.02*	4.5 (0.6)	4.3 (0.9)	.21
Opportunity	4.4 (0.7)	4.1 (1.1)	.08	4.4 (0.7)	4.0 (1.3)	.10	4.4 (0.8)	4.3 (0.7)	.54
General acceptability	4.4 (0.5)	4.3 (1.0)	.56	4.4 (0.5)	4.1 (1.1)	.28	4.4 (0.5)	4.5 (0.8)	.63
Overall score[e]	*4.3 (0.5)*	*4.0 (0.8)*	*.007*	*4.3 (0.5)*	*3.8 (0.9)*	*.008*	4.3 (0.4)	4.2 (0.6)	.41

[a] TFA: theoretical framework for acceptability.

[b] Statistically significant results are given in italic.

[c] TKA: total knee arthroplasty.

[d] THA: total hip arthroplasty.

[e] Burden, ethicality, and opportunity costs were reversed to compute the overall TFA score.

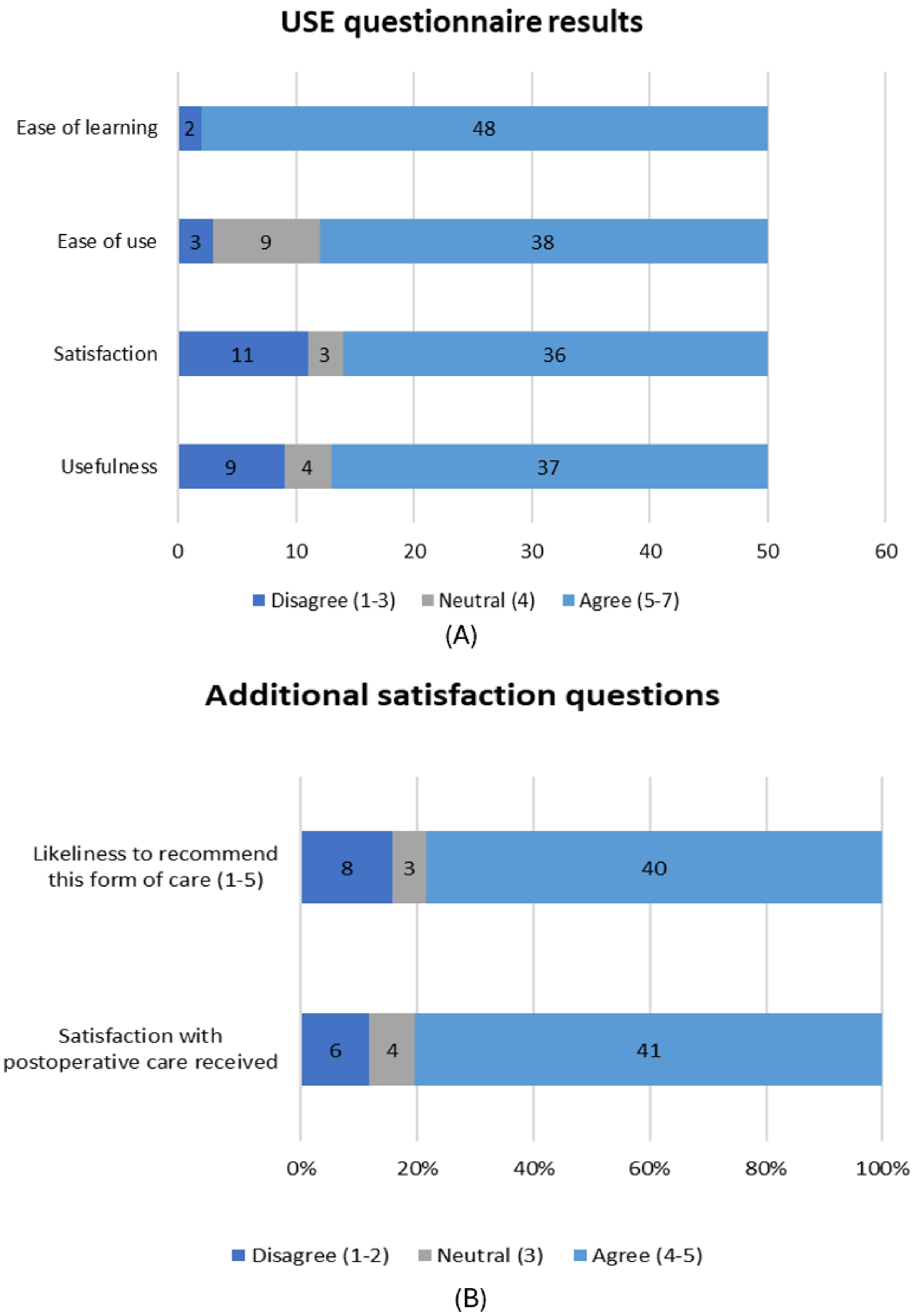
Satisfaction

The USE questionnaire and the additional satisfaction questions revealed good levels of satisfaction with the use of the app (Table 4 and Figure 4). Patients consistently rated the app as useful and easy to integrate into their recovery process.

Table . Usefulness, Satisfaction, and Ease of Use questionnaire results 3 months post operation.

Variable	Participants (n=51)
Usefulness score (1-7)	
Disagree (1-3), n (%)	9 (18)
Neutral (4), n (%)	4 (8)
Agree (5-7), n (%)	37 (74)
Missing, n (%)	1 (1.9)
Mean (SD)	5.0 (1.5)
Satisfaction (1-7)	
Disagree (1-3), n (%)	11 (22)
Neutral (4), n (%)	3 (6)
Agree (5-7), n (%)	36 (72)
Missing, n (%)	1 (1.9)
Mean (SD)	4.9 (1.5)
Ease of use (1-7)	
Disagree (1-3), n (%)	3 (6)
Neutral (4), n (%)	9 (18)
Agree (5-7), n (%)	38 (76)
Missing, n (%)	1 (1.9)
Mean (SD)	5.4 (1.1)
Ease of learning (1-7)	
Disagree (1-3), n (%)	2 (4)
Neutral (4), n (%)	0 (0)
Agree (5-7), n (%)	48 (96)
Missing, n (%)	1 (1.9)
Mean (SD)	5.9 (0.9)
Additional questions	
Level of satisfaction with postoperative care (1-5)	
Disagree (1-2), n (%)	6 (12)
Neutral (3), n (%)	4 (8)
Agree (4-5), n (%)	41 (80)
Mean (SD)	4.1 (1.2)
Likelihood to recommend care to friend or family member (1-5)	
Disagree (1-2), n (%)	8 (16)
Neutral (3), n (%)	3 (6)
Agree (4-5), n (%)	40 (78)
Mean (SD)	4.2 (1.3)

Figure 4. Satisfaction questionnaire results: (A) USE questionnaire results on a stacked bar chart for Likert data; (B) additional satisfaction question results on a 100% stacked bar chart. USE: Usefulness, Satisfaction, and Ease of Use.


USE questionnaire results
Ease of learning
Ease of use
Satisfaction
Usefulness
2
48
3
9
38
11
3
36
9
4
37
0
10
20
30
40
50
60
■ Disagree (1-3)
■ Neutral (4)
■ Agree (5-7)
(A)
Additional satisfaction questions
Likeliness to recommend this form of care (1-5)
Satisfaction with postoperative care received
8
3
40
6
4
41
0%
20%
40%
60%
80%
100%
■ Disagree (1-2)
■ Neutral (3)
■ Agree (4-5)
(B)

Patient-Reported Outcome Measures

There was a statistically significant increase in overall Oxford score and OHS at 31.2 days post operation, but only at 89.4 days post operation in OKS (Table 5 and Figure 5). As for the VAS, there was a significant decrease in overall, hip, and knee VAS scores at 31.2 days post operation.

Table . Oxford scores and VAS[a] results at 1, 3, and 6 months post operation[b].

Sample time point	Preop[c]		First postop (with preop data)[d]				Second postop (with preop data)[e]				Third postop (with preop data)[f]			
	Participants, n	Mean (SD)	Participants, n	Mean (SD)	Mean difference vs pre	*P* value (*t* test)	Participants, n	Mean (SD)	Mean difference vs pre	*P* value (*t* test)	Participants, n	Mean (SD)	Mean difference vs pre	*P* value (*t* test)
Oxford score overall (0 - 48)	49	24.1 (10.1)	49	*30.4 (8.6)*	*6.3*	*<.001*	21	*37.8 (5.7)*	*10.4*	*<.001*	9	*39.1 (3.8)*	*12.2*	*.002*
Oxford Hip Score (0 - 48)	20	22.8 (11.8)	20	*35.0 (6.4)*	*12.3*	*<.001*	8	*40.5 (4.4)*	*13.1*	*.01*	1	—[g]	—	—
Oxford Knee Score (0 - 48)	29	25.0 (8.9)	29	27.2 (8.6)	2.3	.17	13	*36.1 (5.9)*	*8.7*	*.001*	8	*39.9 (3.3)*	*12.5*	*.004*
VAS overall (0 - 10)	49	4.6 (2.6)	49	*2.3 (2.1)*	*−2.3*	*<.001*	22	*0.9 (1.1)*	*−3.1*	*<.001*	9	*1.1 (0.4)*	*−2.4*	*.007*
VAS hip (0 - 10)	20	4.5 (2.6)	20	*1.5 (1.4)*	*−3.0*	*<.001*	9	*0.7 (0.3)*	*−2.7*	*.009*	1	—	—	—
VAS knee (0 - 10)	29	4.7 (2.6)	29	*2.9 (2.2)*	*−1.8*	*<.001*	13	*1.0 (0.4)*	*−3.4*	*<.001*	8	*0.8 (0.3)*	*−2.7*	*.004*

[a] VAS: visual analog scale.

[b] Statistically significant results are given in italic.

[c] Gap with surgery (days): mean 18.8 (SD 10.1, range 3-36).

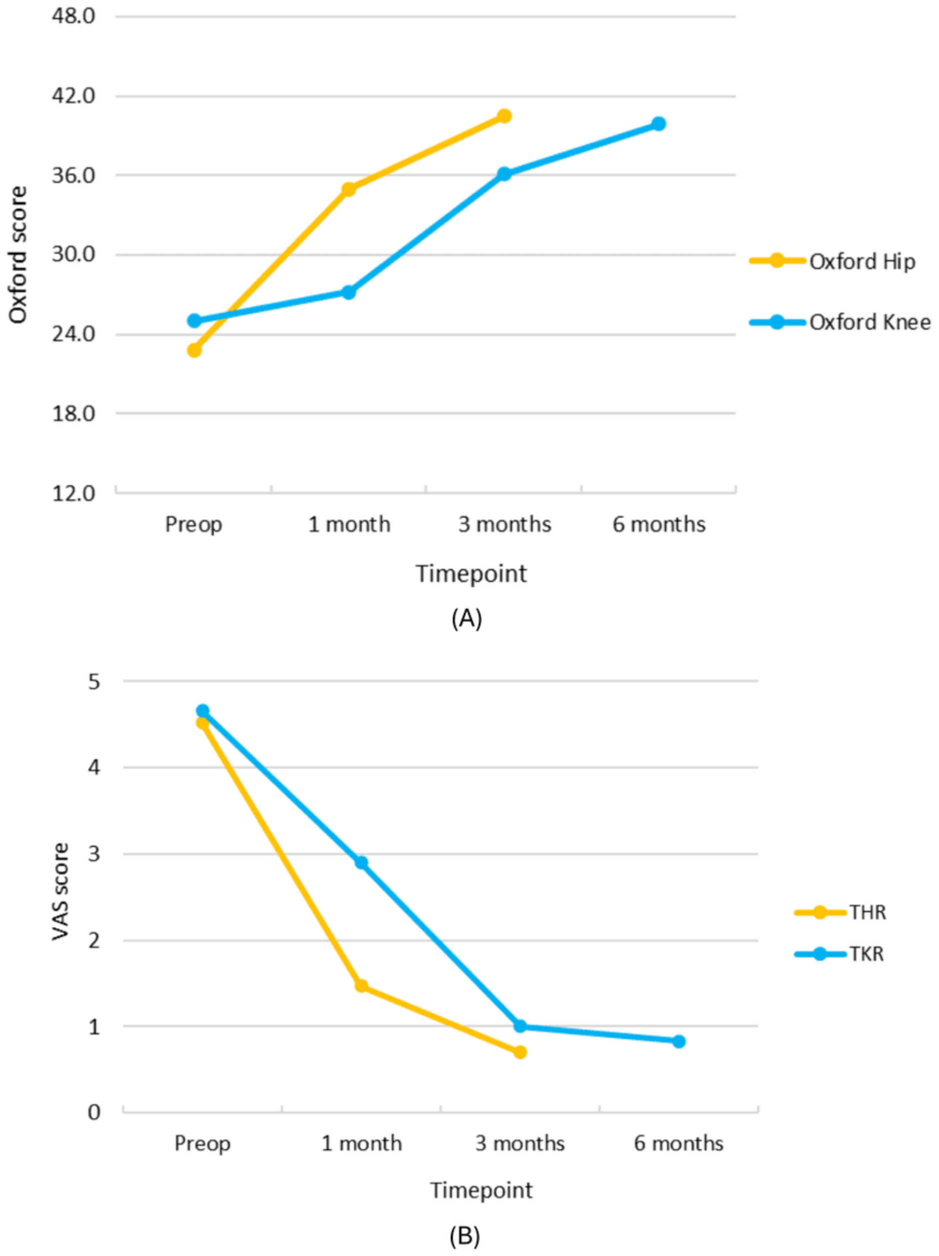
[d] Gap with surgery (days): mean 31.2 (SD 2.8, range 30-44).

[e] Gap with surgery (days): mean 89.4 (SD 5.1, range 67-93).

[f] Gap with surgery (days): mean 180.1 (SD 0.4).

[g] Not available.

Figure 5. Pre- and postoperative (1, 3, and 6 mo) patient-reported outcome measure results by procedure: (A) Oxford scores and (B) VAS scores. THR: total hip replacement; TKR: total knee replacement; VAS: visual analog scale.



App Nondownload Rate

Of the 25 participants who did not download the app, 7 provided a reason. Their responses included technological concerns (n=2), feeling preoccupied and anxious preparing for the procedure (n=1), family issues (n=2), missing download instructions (n=1), and privacy concerns (n=1). The remaining 18 participants were also contacted on multiple occasions but never responded.

Discussion

Principal Findings

In this study, we investigated the acceptability of the use of the mymobility app for the postoperative care of 100 participants undergoing THA and TKA. We sought to investigate acceptability using the TFA to ensure all aspects of acceptability were measured.

Preoperatively, overall acceptability was satisfactory, with all but one TFA item scoring above 4 out of 5. The sole exception was burden (3.9/5), which may reflect lower familiarity with smartphone apps among older participants in our sample. We also found that higher education was associated with a higher level of self-efficacy. While general and overall acceptability were not significantly higher with higher education, this trend aligns with Wang et al’s [15], Nuveo et al’s [31], and Lee et al’s [32] results, who described trends of higher engagement and adherence with higher education and lower receptivity to digital health technologies with lower education. Thus, these findings reveal the importance for app developers to consider this gap and tailor their content to diverse populations, notably through the inclusion of images or videos, which are more easily accessible than heavy-text content.

When comparing pre- and postoperative TFAs, although overall scores were satisfactory at both time points, we found a significant decrease in overall acceptability postoperatively. We suspect this decrease is attributed to elevated expectations toward the app, which were adjusted following its use. However, we believe acceptability could also have been influenced by the natural process of recovery. Indeed, the decline in TFA was significant in TKA, but not in THA. PROM results showed the same trend, with no improvement until 3 months post operation in TKA as opposed to 1 month post operation in THA. This is consistent with Booth et al's [19] and Bourne et al's [33] results showing that TKA patients felt a higher need for in-person rehabilitation as opposed to THA and that short-term outcomes are superior in THA than in TKA. When considering the TFA items that decreased in TKA, which are linked to the perceived helpfulness of the app, this pushes us to believe that the decrease in postoperative acceptability was likely also driven by the surgical outcomes. This illustrates the possibility for app developers to better address the challenges in TKA rehabilitation, notably through adjustment of patient expectations following this procedure.

When looking at participants who did not download the app, reasons included concerns regarding technology in 2 cases, which might be due to the age of the participants (77-79 y). Additionally, 1 participant expressed privacy concerns. Although we explained how data collection and privacy functioned within the app, we still expected this number to be higher. This result is a positive sign for patient trust toward platforms used to provide care, although the need to discuss patient confidentiality within the app remains important.

Comparison With Previous Work

Multiple studies have investigated patient perspectives on the use of mHealth apps for joint arthroplasty. These studies all reported a high level of engagement, adherence, and satisfaction [13,15-19,31]. Factors that influenced the results positively included higher education and supportive environments, whereas those that did negatively included higher age, lower economic status, and worsened physical or psychological condition [15,17,31]. Patients also described better health care accessibility and reduced isolation with these interventions [20]. Our results showed that higher education was associated with greater acceptability. However, unlike prior work, we did not observe other demographic characteristics to significantly influence acceptability.

Strengths and Limitations

The main strength of this study is the measure of acceptability using a validated tool (TFA), which allowed us to assess all domains of this concept. Using this framework, we measured acceptability preoperatively, prior to exposure to the app, and postoperatively, after using the app. This allowed us to interpret how preexposure expectations were met following use of the app. Other strengths include the use of a single app, which reduced heterogeneity; the ITT design, which reflects real-world adherence; and the diverse population.

This study also has limitations. The main limitation is the absence of a control group, which limits the ability to attribute the results specifically to the app. This limitation is a result of ethical concerns brought by our institutional review board, which prohibited us from having a control group that could be denied a beneficial intervention. We also observed a high rate of unfilled questionnaires, which could have created a response bias. We believe this is due to most questionnaires being distributed by email or phone, creating a limiting step in communication. Other important limitations include the use of a single app, limiting the generalizability of the results; a potential selection bias toward patients willing to use the app; and statistical analysis being only performed with *t* tests, which might have limited the power of the conclusions. In addition, there was no assessment of baseline technological literacy, which could have affected results despite the app's technological guidance features.

Conclusion

There was a good level of acceptability with the use of mymobility for the postoperative management of THA and TKA, although acceptability decreased postoperatively. This decrease could signify high expectations toward the app preoperatively or higher than expected difficulty and pain in the early postoperative period. Higher education was associated with higher preoperative acceptability. TKA was associated with lower postoperative acceptability, which could be related to the delayed recovery in TKA when compared to THA. These 2 trends have previously been described in the literature, indicating potential gaps to address for app developers. We found a good level of satisfaction with the app, and Oxford scores showed better outcomes in THA than TKA, whereas VAS showed equally significant decreases in pain with both procedures. Privacy concerns were rare, with only 1 patient reporting them.

The impact of the natural recovery process in total joint arthroplasty on the acceptability of mHealth remains unclear. This gap could be addressed through a comparison of acceptability pre- and postoperatively in cohorts using the app and cohorts with standard care.

Acknowledgments

The authors sincerely thank St. Mary's Hospital for its support. The authors declare that no generative artificial intelligence was used in the preparation and writing of this article. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of the work.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

All data generated or analyzed during this study are included in this published article (and the multimedia appendices).

Conflicts of Interest

A Albers is a paid consultant for Zimmer-Biomet. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Overview of data collected by the research team and by the mymobility app.
[DOCX File, 16 KB - humanfactors_v13i1e79682_app1.docx]

Multimedia Appendix 2
Overview of the mymobility functions and interface. (A) Features of the app with depictions of the app interface and (B) Walk-AI, the mymobility artificial intelligence–assisted gait speed analysis predicting and determining level of recovery based on gait speed.
[DOCX File, 640 KB - humanfactors_v13i1e79682_app2.docx]

Multimedia Appendix 3
Modified theoretical framework for acceptability questionnaire.
[DOC File, 51 KB - humanfactors_v13i1e79682_app3.doc]

Multimedia Appendix 4
Usefulness, Satisfaction, and Ease of Use questionnaire.
[DOC File, 37 KB - humanfactors_v13i1e79682_app4.doc]

Multimedia Appendix 5
Additional satisfaction questions.
[DOC File, 33 KB - humanfactors_v13i1e79682_app5.doc]

Multimedia Appendix 6
Oxford Hip and Oxford Knee Scores.
[DOC File, 82 KB - humanfactors_v13i1e79682_app6.doc]

Multimedia Appendix 7
Visual analog scale for pain.
[DOC File, 268 KB - humanfactors_v13i1e79682_app7.doc]

References

1. Hospital stays in Canada, 2021–2022. : Canadian Institute for Health Information; 2023 Feb 23 URL: https://bit.ly/3N6xCNM [accessed 2025-12-09]
2. Shichman I, Roof M, Askew N, et al. Projections and epidemiology of primary hip and knee arthroplasty in Medicare patients to 2040-2060. JB JS Open Access 2023;8(1):e22.00112. [doi: 10.2106/JBJS.OA.22.00112] [Medline: 36864906]
3. Aging and chronic diseases: a profile of Canadian seniors. : Public Health Agency of Canada; 2020 URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/aging-chronic-diseases-profile-canadian-seniors-report.html [accessed 2025-12-09]
4. Sherif A. Global smartphone penetration rate as share of population from 2016 to 2024. Statista. 2025. URL: https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/ [accessed 2025-12-09]
5. Wells C, Spry C. An overview of smartphone apps: CADTH horizon scan. : Ottawa (ON): Canadian Agency for Drugs and Technologies in Health; 2022 URL: https://www.ncbi.nlm.nih.gov/books/NBK595384 [accessed 2025-12-09]
6. Andrawis JP, Muzykewicz DA, Franko OI. Mobile device trends in orthopedic surgery: rapid change and future implications. Orthopedics 2016;39(1):e51-e56. [doi: 10.3928/01477447-20151228-01] [Medline: 26730684]
7. Dionisi S, Giannetta N, Di Simone E, et al. The use of mHealth in orthopedic surgery: a scoping review. Int J Environ Res Public Health 2021 Nov 28;18(23):12549. [doi: 10.3390/ijerph182312549] [Medline: 34886274]

8. Hayat J, Ramadhan M, Gonnah AR, Alfadhli A, Al-Naseem AO. The role of mobile health technology in perioperative spinal care: a systematic scoping review and narrative synthesis. Cureus 2024 Feb;16(2):e54254. [doi: 10.7759/cureus.54254] [Medline: 38496189]
9. Mapinduzi J, Ndacayisaba G, Verbrugghe J, Timmermans A, Kossi O, Bonnechère B. Effectiveness of mHealth interventions to improve pain intensity and functional disability in individuals with hip or knee osteoarthritis: a systematic review and meta-analysis. Arch Phys Med Rehabil 2025 Feb;106(2):280-291. [doi: 10.1016/j.apmr.2024.06.008] [Medline: 38945508]
10. Pritwani S, Shrivastava P, Pandey S, et al. Mobile and computer-based applications for rehabilitation monitoring and self-management after knee arthroplasty: scoping review. JMIR Mhealth Uhealth 2024 Jan 26;12:e47843. [doi: 10.2196/47843] [Medline: 38277195]
11. Dieter V, Janssen P, Krauss I. Efficacy of the mHealth-based exercise intervention re.flex for patients with knee osteoarthritis: pilot randomized controlled trial. JMIR Mhealth Uhealth 2024 Sep 9;12:e54356. [doi: 10.2196/54356] [Medline: 39250181]
12. Özden F, Sarı Z. The effect of mobile application-based rehabilitation in patients with total knee arthroplasty: a systematic review and meta-analysis. Arch Gerontol Geriatr 2023 Oct;113:105058. [doi: 10.1016/j.archger.2023.105058] [Medline: 37172329]
13. Monárrez R, Mohamadi A, Drew JM, Abdeen A. Mobile application’s effect on patient satisfaction and compliance in total joint arthroplasty: a systematic review and meta-analysis. J Am Acad Orthop Surg Glob Res Rev 2023 Sep 1;7(9):e22.00200. [doi: 10.5435/JAAOSGlobal-D-22-00200] [Medline: 37678829]
14. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
15. Wang Q, Lee RLT, Hunter S, Zhu A, Chan SWC. Patient engagement in a mobile app-based rehabilitation program for total hip or knee arthroplasty: secondary data analysis of a randomized controlled trial. JMIR mHealth uHealth 2024 Oct 1;12:e57635. [doi: 10.2196/57635] [Medline: 39353187]
16. Hong M, Loeb J, Yang M, Bailey JF. Postoperative outcomes of a digital rehabilitation program after total knee arthroplasty: retrospective, observational feasibility study. JMIR Form Res 2022 Sep 19;6(9):e40703. [doi: 10.2196/40703] [Medline: 36121690]
17. Wang Q, Gao J, Zhu A, Cheng C, Chan SWC. Patient engagement in mobile technology-based rehabilitation for arthroplasty: a scoping review. J Clin Nurs 2025 Sep;34(9):3489-3503. [doi: 10.1111/jocn.17802] [Medline: 40346770]
18. Ramkumar PN, Haeberle HS, Ramanathan D, et al. Remote patient monitoring using mobile health for total knee arthroplasty: validation of a wearable and machine learning-based surveillance platform. J Arthroplasty 2019 Oct;34(10):2253-2259. [doi: 10.1016/j.arth.2019.05.021] [Medline: 31128890]
19. Booth MW, Riegler V, King JS, Barrack RL, Hannon CP. Patients’ perceptions of remote monitoring and app-based rehabilitation programs: a comparison of total hip and knee arthroplasty. J Arthroplasty 2023 Jul;38(7S):S39-S43. [doi: 10.1016/j.arth.2023.04.032] [Medline: 37088222]
20. Wang Q, Lee RLT, Hunter S, Chan SWC. Patients’ experiences of using a mobile application-based rehabilitation programme after total hip or knee arthroplasty: a qualitative descriptive study. BMC Nurs 2023 Jul 27;22(1):246. [doi: 10.1186/s12912-023-01409-3] [Medline: 37496003]
21. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480. [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
22. Mymobility® with Apple Watch. Zimmer Biomet. 2023. URL: https://www.zimmerbiomet.com/en/products-and-solutions/zb-edge/mymobility.html#01-overview [accessed 2025-12-09]
23. Mymobility® Patient App. Google Play. URL: https://play.google.com/store/apps/details?id=com.mymobility [accessed 2025-12-09]
24. Privacy notice. Zimmer Biomet. 2023. URL: https://www.zimmerbiomet.com/en/corporate/privacy-notice/privacy-notice-english.html [accessed 2025-12-09]
25. PIPEDA requirements in brief. Office of the Privacy Commissioner of Canada. URL: https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/ [accessed 2025-12-09]
26. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022 Mar 1;22(1):279. [doi: 10.1186/s12913-022-07577-3] [Medline: 35232455]
27. Lund AM. Measuring usability with the USE questionnaire. Usability Interface 2001;8(2):3-6 [FREE Full text]
28. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996 Mar;78(2):185-190. [doi: 10.1302/0301-620X.78B2.0780185] [Medline: 8666621]
29. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br 1998 Jan;80-B(1):63-69. [doi: 10.1302/0301-620X.80B1.0800063]
30. Kersten P, White PJ, Tennant A. Is the pain visual analogue scale linear and responsive to change? An exploration using Rasch analysis. PLoS ONE 2014;9(6):e99485. [doi: 10.1371/journal.pone.0099485] [Medline: 24921952]

31. Nuevo M, Modrego A, Rodríguez-Rodríguez D, et al. The influence of sociodemographic and health factors on adherence to home-based rehabilitation after fast-track total knee arthroplasty: secondary analysis of a randomized controlled trial. Disabil Rehabil 2025 Sep;47(18):4728-4737. [doi: 10.1080/09638288.2025.2458191] [Medline: 39879639]
32. Lee CMY, Thomas E, Norman R, et al. Educational attainment and willingness to use technology for health and to share health information - the reimagining healthcare survey. Int J Med Inform 2022 Aug;164:104803. [doi: 10.1016/j.ijmedinf.2022.104803] [Medline: 35644052]
33. Bourne RB, Chesworth B, Davis A, Mahomed N, Charron K. Comparing patient outcomes after THA and TKA: is there a difference? Clin Orthop Relat Res 2010 Feb;468(2):542-546. [doi: 10.1007/s11999-009-1046-9] [Medline: 19760472]

Abbreviations

mHealth: mobile health
OHS: Oxford Hip Score
OKS: Oxford Knee Score
PROM: patient-reported outcome measure
TFA: theoretical framework for acceptability
THA: total hip arthroplasty
TKA: total knee arthroplasty
USE: Usefulness, Satisfaction, and Ease of Use
VAS: visual analog scale

Edited by A Kushniruk; submitted 29.Jun.2025; peer-reviewed by W Glinkowski, X Zhang; revised version received 23.Oct.2025; accepted 20.Nov.2025; published 27.Jan.2026.

Please cite as:
Louni Y, Laroche M, Alnasser A, Abuhaneya M, Belzile E, Baskaran S, Mutch J, Albers A
Acceptability of a Digital Care App in Patients Undergoing Hip and Knee Arthroplasty: Prospective Cohort Study
JMIR Hum Factors 2026;13:e79682
URL: https://humanfactors.jmir.org/2026/1/e79682
doi:10.2196/79682

Long-Term Effects of Multimedia Education and “Foot Alerts” Through the MyU App on Diabetic Foot Care in Individuals With Moderate-High Risk: Pilot Randomized Controlled Trial

Ameenah Alawadhi[1], BSc, MSc, PhD; Kay Scarsbrook Khan[2,3], BSc; Grace Messenger[2,4], BSc; Alice Cusworth[2], BSc; Mohammad Assi[5], BSc; Stuart R Gray[6], BSc, MSc, PhD; Ebaa Al-Ozairi[1], BMedSci, MBChB, MRCPE, CCT, FACN, FACE

[1]Clinical Care Research and Trials Unit, Dasman Diabetes Institute, P.O. Box 1180, Dasman 15462, Kuwait City, Kuwait
[2]Podiatry Unit, Dasman Diabetes Institute, Kuwait City, Kuwait
[3]Podiatry Unit, Harrogate and District NHS Foundation Trust, North Yorkshire, United Kingdom
[4]Podiatry Unit, North Cumbria Integrated Care NHS Foundation Trust, Cumbria, United Kingdom
[5]Technology Department, Dasman Diabetes Institute, Kuwait City, Kuwait
[6]School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, United Kingdom

Corresponding Author:
Ebaa Al-Ozairi, BMedSci, MBChB, MRCPE, CCT, FACN, FACE
Clinical Care Research and Trials Unit, Dasman Diabetes Institute, P.O. Box 1180, Dasman 15462, Kuwait City, Kuwait

Abstract

Background: Diabetic foot ulceration (DFU) is the leading cause of nontraumatic amputations in people with diabetes. Research shows that improving patient awareness can result in short-term improvements, but Cochrane reviews report insufficient high-quality evidence.

Objective: This study aims to investigate the effects of multimedia presentation and smartphone alerts to enhance long-term knowledge and foot care behaviors in individuals at moderate-to-high risk of DFU.

Methods: Participants were randomized to a control group, receiving usual diabetic foot care advice (n=40), or an intervention group, receiving a multimedia diabetic foot care presentation and regular “foot alerts” through the MyU smartphone app on top of usual care (n=37). Patient’s knowledge and behaviors related to diabetic foot care were assessed at baseline and after 12 months. Repeated measures ANOVA was conducted in both intention-to-treat and per-protocol analyses to evaluate the intervention’s effectiveness.

Results: The findings were consistent across intention-to-treat and per-protocol analyses. In the intervention group, the number of podiatry visits was positively correlated with improved foot care behavior (r=0.408; P=.02), while the control group showed a negative correlation (r=−0.402; P=.02). No significant correlations were observed with knowledge scores. Although no significant time×group interactions were seen, the main effects of time were found for both knowledge (η^2=0.12; P=.004) and behavior scores (η^2=0.31; P<.001). Post hoc analysis showed a decline in knowledge scores in the control group (Cohen d=−0.24; P=.007) and improvements in behavior scores in both groups (Cohen d: intervention=0.61, control=0.63; all P<.001).

Conclusions: The MyU app−based multimedia intervention was associated with improved foot care behaviors over 12 months, indicating potential benefits as an adjunct to usual diabetic foot care. However, no significant changes in diabetic foot care knowledge were observed. These findings suggest that while the applied digital multimedia tool may support behavior change, further research is needed to enhance knowledge retention and clinical impact. The study revealed that multimedia education alone may not be effective for long-term improvement in foot self-care knowledge and behavior among individuals at moderate-high DFU risk, but the reinforcement of educational material during follow-up podiatry visits could be effective.

Trial Registration: ClinicalTrials.gov NCT03934944; https://www.clinicaltrials.gov/study/NCT03934944

(JMIR Hum Factors 2026;13:e78261) doi:10.2196/78261

KEYWORDS

behavior; DFU; diabetic foot ulceration; foot care; multimedia education; smartphone

Introduction

The escalating incidence of diabetes worldwide has turned it into a global pandemic impacting over 500 million individuals [1]. Health costs associated with diabetes are expected to reach US $1054 billion by 2045, up from US $966 billion in 2021 [2], with 80% of this spent on managing avoidable complications [3]. One of the most common long-term complications of diabetes is diabetic foot ulceration (DFU), with a lifetime risk of developing a DFU of between 12% and 25% [4-6]. DFU continues to be the leading cause of nontraumatic lower limb amputation [7], accounting for 84% of lower limb amputations [8]. Studies have shown that people with diabetes are 10 - 30 times more likely to have a lower limb amputation than someone without diabetes [9,10]. Furthermore, Armstrong et al [11] reported a 5-year mortality rate in people with DFU of as much as 43% at 1 year, worse than some of the most common cancers, and postamputation rates have shown that only 50% of patients survive the first year following amputation [12].

The early detection and treatment of minor foot injuries have been shown to reduce amputation rates by between 49% and 85% [13], yet amputation rates continue to rise [14]. Improved patient awareness and correct foot self-care practices can interrupt the amputation pathway [6]. Although systematic reviews have demonstrated the efficacy of foot care education in enhancing self-care behaviors and knowledge [15-17], there are insufficient data to support its clinical use in reducing ulceration and amputation incidence [18,19]. Furthermore, the delivery of education sessions is particularly difficult for those with limited literacy, and debates are ongoing about whether video-based or printed materials are most effective [20,21]. Comprehensive foot care education, which encompasses assessment, discussion, counseling, home visits, and telephone calls, has been proposed to enhance diabetic foot care and educational programs [18]. However, a systematic review of 6 randomized controlled trials (RCTs) revealed insufficient evidence to support its efficacy [22]. More recently, it has been suggested that mobile apps can effectively prevent DFU recurrence [23] and improve diabetic foot care outcomes by incorporating at least 1 information communication tool [24]. However, many of the RCTs focused on patients at low risk, limiting their utility to support improved clinical outcome in the measure of reducing the rate of ulceration [18,19,22].

Given these findings and the estimated financial savings in the prevention of DFU, there is a need to focus on its prevention [25]. This study, therefore, aims to (1) establish if the use of a multimedia presentation and weekly alerts from a smartphone app as an education tool improves long-term knowledge and foot care behaviors of patients with diabetes when compared to usual care, (2) assess the association between the change in participants' knowledge and behavior and the number of podiatry appointments over the study period, and (3) compare the number of new DFUs and hemoglobin A_{1c} (HbA_{1c}) levels between the 2 study groups.

Methods

Study Design

This randomized, investigator-blinded, 2-armed, pilot study was conducted on patients with diabetes who visited Dasman Diabetes Institute (Kuwait City, Kuwait) clinics from January 2019 to May 2024. Each participant was fully informed about the study prior to giving their written informed consent.

Eligibility Criteria

Participants had to meet the following inclusion criteria: (1) adult patients (≥18 years of age) with type 2 diabetes or adults with type 1 diabetes for ≥5 years, (2) be of medium or high DFU risk, defined as having at least 1 diabetic foot risk factor (loss of pain perception, peripheral vascular disease, foot deformity, history of DFU, or amputation), (3) able to understand study procedures and comply with them for the entire length of the study, and (4) must own a smartphone with internet access and agree to have the phone app uploaded for the duration of the study. People with chronic kidney disease, cognitive impairment, active psychiatric illness, inability to give written informed consent, hearing or visual impairment, or phone app inaccessibility for more than 4 weeks were excluded from the study.

Procedures

Overview

The study consisted of 2 main visits (at baseline and 12 mo) and up to 6 interim follow-up visits during the 1-year study period. A total of 98 participants were randomized by an independent researcher to either the control ("Usual Care" section) or intervention ("Multimedia Educational Intervention" section) arm in a 1:1 ratio. Participants were instructed not to disclose any information about their assigned group to the research team to maintain confidentiality. All written information for this study went through a process of back translation from English to Arabic, focusing on conceptual rather than literal translations and using natural language for a broad audience [26]. Prior to the intervention period, participants underwent diabetic foot assessment and were classified based on their diabetic foot risk as per NICE (National Institute for Health and Care Excellence) guidelines [27], with participants with moderate risk attending 2 additional follow-up visits at 6-month intervals and participants with high risk attending 6 additional follow-up visits at 2-month intervals. At the baseline visit, HbA_{1c} was measured, and participants received their assigned diabetic foot care education route. Then, they were asked to complete a diabetic foot care knowledge questionnaire and a foot care and diabetes self-care behaviors questionnaire. Interim follow-up visits included a visual inspection of feet, verbal advice in the usual manner (for both groups), and an educational audio-visual prerecorded presentation (for intervention group only). At the 12-month visit, all participants completed a foot care knowledge questionnaire, a foot care and diabetes self-care behaviors questionnaire, and underwent an HbA_{1c} test.

Usual Care

The control group received usual care and the usual route of education with educational leaflets to take home. Participants received verbal advice during their podiatry appointment about daily foot inspections, wearing well-fitting shoes, what to do in an emergency, and the use of emollient. More specific advice was given depending on the participant's risk classification. This usual route involved information by the podiatrist translated by a bilingual but not native Arabic-speaking nurse.

Multimedia Educational Intervention

The intervention group received an educational presentation and a smartphone app providing weekly foot alerts for the duration of the study, in addition to the usual care, education route, and leaflet. The details of the MyU app interface, including the layout and key features used to deliver educational content, are illustrated in Figure 1. The educational presentation was a 10-minute audio-visual prerecorded presentation based on the internationally recognized advice on preventing diabetic foot complications including good diabetes control, daily foot inspections, and footwear advice [27-29]. The research podiatry clinic allowed only 1 participant to attend at a time, ensuring privacy for viewing the educational video. In addition, a teaching- or education-based smartphone app called MyU was installed onto their smartphones, allowing educational information to be uploaded and delivered remotely. The app also remotely recorded the number of times the content was accessed by each participant. Adherence to the intervention was monitored based on these access logs, where each "access" was defined as opening the app and viewing the educational content. However, more detailed engagement metrics such as video completion or time spent on content were not captured. Adherence (%) was calculated as the percentage of scheduled videos viewed, based on the total video views recorded by the MyU app for each participant. Both the absolute number of videos viewed (count) and the corresponding percentage adherence were reported. This app also delivered a weekly "foot-alert" (notification) to remind the participant to view the uploaded educational video. There were 12 unique educational videos "alerts" repeated at 3-month cycles. Repeating the educational method has been shown to improve health literacy [30]. All foot alerts were uploaded in both Arabic and English.

Figure 1. MyU app interface showing the layout and key features used to deliver educational content. The home screen provides access to 12 educational videos, which were assigned to participants according to their recruitment date. Weekly push notifications (shown at the bottom of the screen) prompted users to engage with the app and view the assigned videos displayed on the Newsfeed page.

Foot Care Knowledge Questionnaire

The foot care knowledge questionnaire used in this study has been adapted from Pollock et al [31] and Rheeder et al [32]. The questionnaire covers 2 sections. The first section consisted of 7 questions with 3 options each: correct, incorrect, and don't know. A correct answer was awarded 1 point, while incorrect or "don't know" responses received 0 points. The second section consisted of 5 questions with multiple-choice options. A correct answer was awarded 1 point, while an incorrect answer received 0 points. The maximal total score is 12 points, with higher scores indicating a better understanding of foot care knowledge.

Foot Care and Diabetes Self-Care Behaviors Questionnaire

The foot care and diabetes self-care behaviors questionnaire has been adopted from the Summary of Diabetes Self-Care Activities (SDSCA). The SDSCA has been found to have both reliability and validity as a standard measure of diabetes self-management [33]. Furthermore, this test has been shown to be reliable and valid, giving consistent results when translated into Arabic in a sample size of 243 participants [34]. The SDSCA scale measures the frequency of each self-care activity in the last 7 days for 4 aspects related to diabetes routine: foot care (8 scaled questions), blood-glucose testing (2 scaled questions), medications adherence (2 scaled questions), and smoking habits (2 questions). For this analysis, the score of the scaled questions (from 0 to 7 d) was calculated as the average of responses within each section [33]. A higher total score indicates better foot self-care behavior.

Sample Size

Although this is a pilot study, a sample size of 98 with a 1:1 ratio between intervention and control group (49:49) was calculated using G*power software (version 3.9.1.2; Heinrich-Heine-Universität Düsseldorf). An a priori difference between 2 independent means tests was performed to find a significance interaction with a power of 80% and probability of type 1 error of 0.05. The effect from a study by Baba et al [35] was a significant mean change of 1.8 (SD 2.6) reduction in foot score (based on the presence and severity of 15 podiatry disorders; n=78) compared to a 0.1 (SD 2.6) reduction in the foot score after receiving interactive education (n=76) over 3 months (effect size=0.58).

Statistical Analysis

Data analysis was performed using SPSS Statistics (version 29.0; IBM Corp). Descriptive statistics were represented as mean (SD) for continuous variables and as frequencies for categorical variables. The Shapiro-Wilk test was used to assess normality for each continuous variable to determine the appropriate use of parametric or nonparametric tests. Chi-square tests and independent 2-tailed *t* tests were conducted to test to examine baseline differences between the groups and to compare the frequency of new DFU cases between the study groups. The Wilcoxon signed-rank test was used when the assumption of normality was violated.

Both intention-to-treat (ITT) and per-protocol (PP) analyses were conducted. The ITT analysis included all randomized participants with the available outcome data, regardless of adherence or dropout status. A linear mixed model was used to evaluate group, time, and group×time interaction effects on the outcome measures: foot care knowledge scores, foot care, blood-glucose testing, and medication adherence behaviors.

Post hoc pairwise comparisons were conducted using the matched pair 2-tailed Student *t* test or Wilcoxon signed-rank test, with Bonferroni adjustment applied for multiple comparisons. Bivariate Pearson or Spearman correlation analyses were performed to evaluate the strength and direction of relationships between the change in foot care knowledge scores and foot self-care behavior: scores and the number of follow-up visits. Changes in scores were calculated as 12 month values − baseline values. *P* values <.05 were considered statistically significant.

Ethical Considerations

Ethical approval was granted by the Dasman Diabetes Institute Ethical Review Committee (RA HM-2018-044). The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines and was registered at ClinicalTrials.gov (NCT03934944). The trial was conducted and reported in accordance with the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) 2011 guidelines (Checklist 1).

Written informed consent was obtained from all participants prior to enrollment. Participant privacy and confidentiality were strictly maintained throughout the study: all data were deidentified and stored on secure, password-protected servers accessible only to authorized research personnel, and no personally identifiable information was included in the analysis or reported in the manuscript. Participants did not receive financial compensation for their participation.

Results

A total of 66 participants completed the study visits and were included in the final PP analysis (Figure 2). Additionally, the ITT analysis included 77 participants who were randomized and had at least 1 outcome measure available. Table 1 shows the baseline characteristics of study participants in both ITT and PP populations. In the ITT population, the mean age of the participants was 60 (SD 7.85) years in the intervention group and 62 (SD 9.31) years in the control group. In both groups, the majority of the study participants were male participants, had attained an undergraduate degree, were nonsmokers, and had type 2 diabetes for over 20 years. In the intervention group, 75.7% (n=28) were classified as having a high DFU risk compared to a 62.5% (n=20) in the control group. All baseline characteristics were statistically comparable between the groups in both ITT and PP analyses, except for BMI, which showed a statistically significant difference between the groups in the ITT sample (*P*=.01) but not in the PP sample. Dropout analysis revealed that 8 participants from the control group (6 with high DFU risk and 2 with moderate risk) and 3 participants from the intervention group (all high risk) discontinued the study.

Figure 2. Participants flow chart. DFU: diabetic foot ulceration; ITT: intention-to-treat; PP: per-protocol.

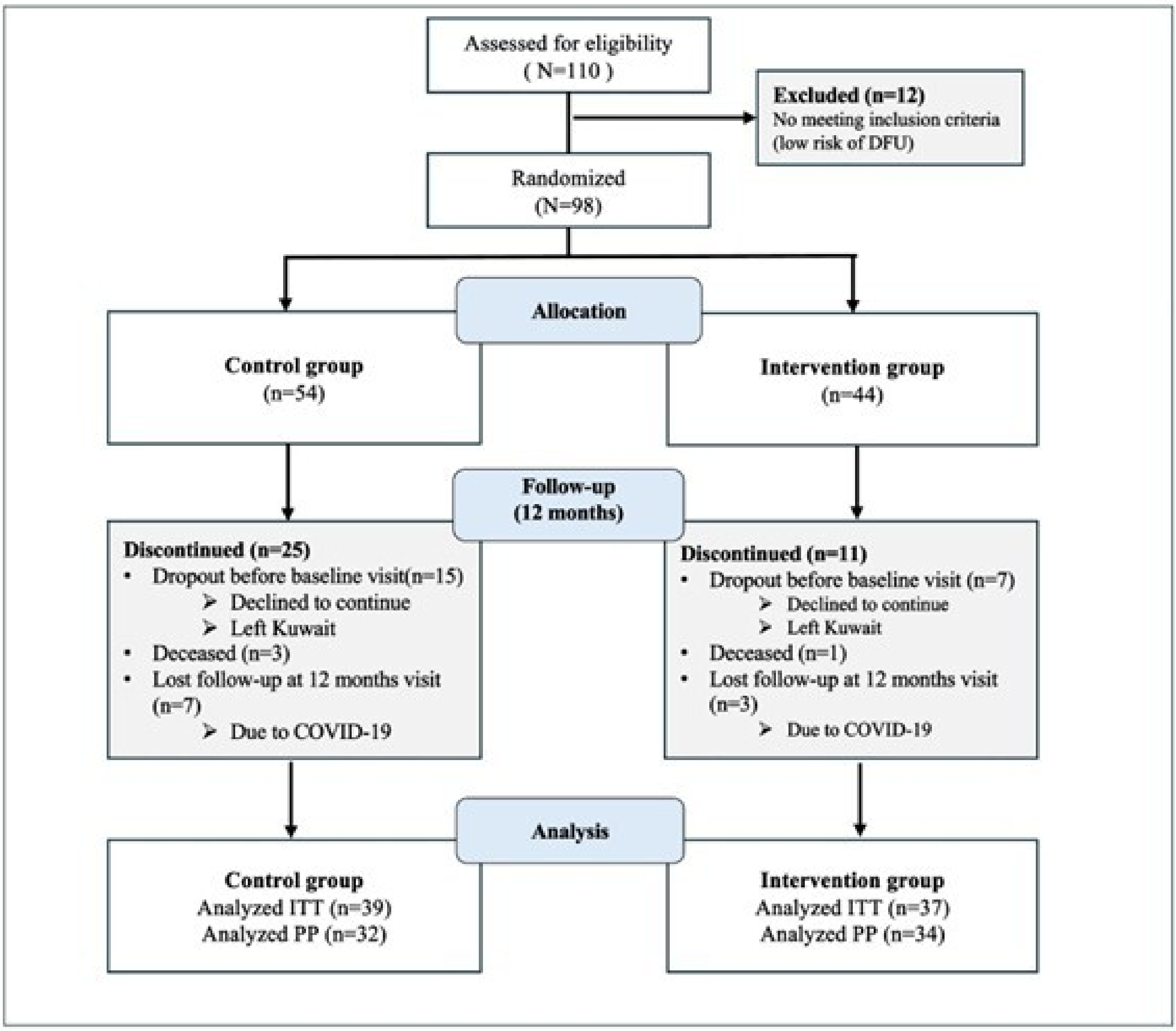


Table . General characteristics according to study group and analysis carried out.

Variable	ITT[a]			PP[b]		
	Intervention (n=37)	Control (n=40)	*P* value	Intervention (n=34)	Control (n=32)	*P* value
Age (y), mean (SD)	60.43 (7.85)	62.39 (9.81)	.34	61.06 (7.87)	62.70 (9.41)	.44
BMI (kg/m^2), mean (SD)	34.05 (5.69)	36.94 (3.2)	.01[c]	34.05 (5.69)	32.09 (6.04)	.18
Sex, n (%)			.83			.65
Male	24 (65)	25 (63)		21 (62)	20 (63)	
Female	13 (35)	15 (38)		13 (38)	12 (38)	
Ethnicity, n (%)			.57			.49
Arabic	2 (5)	3 (8)		2 (6)	3 (9)	
Asian	0 (0)	1 (3)		0 (0)	1 (3)	
Kuwaiti	35 (95)	36 (90)		32 (94)	28 (88)	
Education level, n (%)			.23			.21
Junior school	8 (22)	2 (5)		6 (18)	1 (3)	
Senior school	9 (24)	11 (28)		9 (27)	7 (23)	
Undergraduate	14 (38)	22 (55)		14 (41)	19 (61)	
Postgraduate	7 (19)	4 (10)		5 (15)	4 (13)	
Smoking status, n (%)			.51			.21
Smoker	6 (16)	5 (13)		6 (18)	3 (9)	
Nonsmoker	26 (65)	23 (62)		26 (81)	21 (62)	
Ex-smoker	7 (21)	5 (13)		7 (21)	3 (9)	
Type of diabetes, n (%)			.58			.33
Type 1	3 (8)	2 (5)		3 (9)	1 (3)	
Type 2	34 (92)	38 (95)		31 (97)	31 (91)	
Duration of diabetes (y), n (%)			.82			.83
0 - 9	4 (11)	5 (13)		4 (13)	3 (9)	
10 - 19	11 (30)	14 (35)		9 (28)	11 (34)	
>20	22 (60)	21 (53)		19 (59)	18 (56)	
HbA$_{1c}$[d], mean (SD)	8.32 (1.87)	8.14 (1.61)	.65	8.29 (2.04)	8.17 (1.68)	.61
Overall ABPI[e], mean (SD)	1.17 (0.13)	1.18 (0.07)	.95	1.186 (0.21)	1.13 (0.14)	.56
Average VPT[f], mean (SD)	52.60 (118.3)	50.86 (113.7)	.94	53.43 (123.55)	55.84 (129.16)	.94
Risk factors, n (%)						
PAD[g]	7 (19)	6 (15)	.57	6 (18)	5 (16)	.82
LOPS[h]	33 (89)	37 (93)	.41	30 (88)	30 (92)	.43
Foot deformity	22 (60)	19 (48)	.13	21 (62)	16 (50)	.33
History of DFU[i]	18 (49)	16 (40)	.5	16 (47)	13 (41)	.59
History of amputation	7 (19)	8 (20)	.61	7 (21)	7 (22)	.89
History of Charcot	5 (14)	4 (10)	.56	4 (12)	3 (9)	.75
DFU risk classification, n (%)			.12			.13

Variable	ITT[a]		PP[b]	
High	28 (76)	25 (63)	25 (76)	18 (58)
Moderate	9 (24)	15 (38)	9 (24)	14 (42)

[a] ITT: intention-to-treat.

[b] PP: per protocol.

[c] *P value <.05

[d] HbA$_{1c}$: hemoglobin A$_{1c}$.

[e] ABPI: ankle-brachial pressure index.

[f] VPT: vibration perception threshold.

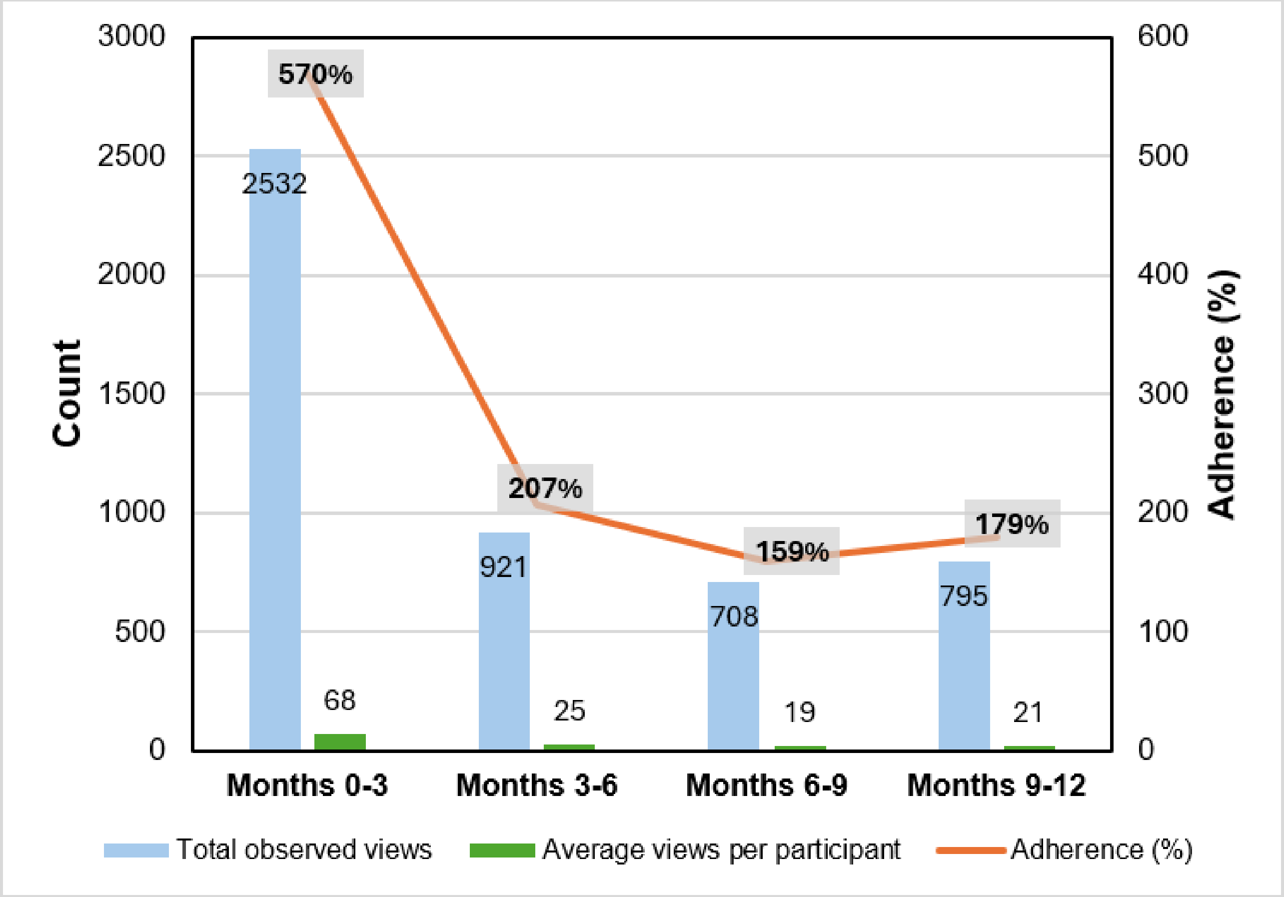
[g] PAD: peripheral artery disease.

[h] LOPS: diabetic peripheral neuropathy with loss of protective sensation.

[i] DFU: diabetic foot ulceration.

Adherence was estimated using the total video views recorded by the MyU app. The actual views exceeded the maximum expected views per interval (444 views), especially at baseline (mo 0 - 3: 2532 views, 570%), indicating repeated viewing. Engagement declined over time (mo 9 - 12: 795 views, 179%). These counts serve as a group-level proxy for adherence (Figure 3).

Figure 3. Total video views (blue bars), average views per participant (green bars), and adherence percentage (line) across study time points, based on MyU app data. Values represent group-level engagement with 12 videos per interval. Adherence (%) represents the percentage of the scheduled videos viewed, calculated from the total video views recorded by the MyU app. Both the absolute number of videos viewed (count) and the corresponding percentage are shown.



Both ITT and PP analyses revealed no significant time×group interactions for diabetic foot care knowledge and behavior scores. However, there were significant main effects of time for both outcomes (ITT: $F_{1,70}$=6.98, P=.01, partial η^2=0.09 for knowledge; $F_{1,70}$=28.36, P<.001, partial η^2=0.29 for behavior; PP: $F_{1,64}$=8.760; P=.004, partial η^2=0.12 for knowledge scores; $F_{1,64}$=28.139; P<.001, partial η^2=0.31 for behavior scores). Post hoc within-group comparisons showed a significant reduction in diabetic foot care knowledge scores (ITT: mean difference −0.878, SE 0.366 points, P value=.02, Cohen d=−0.21; PP: mean difference −1, SE 0.357 points, P value=.007, Cohen d=−0.22) in the control group, with no change in the intervention group (ITT: mean difference −0.45, SE 0.344

points, P=.2, Cohen d=−0.38; PP: mean difference −0.47, SE 0.346 points, $P_{t\text{ test}}$=.18, Cohen d=−0.48). For diabetic foot care behavior, post hoc within-group comparisons showed a significant increase in both control (ITT: mean difference 1.129, SE 0.29 points, P<.001, Cohen d=0.62*;* PP: mean difference 1.109, SE 0.3 points, P<.001, Cohen d=0.61) and intervention (ITT: mean difference 1.139, SE 0.3 points, P<.001, Cohen d=0.60; PP: mean difference 1.121, SE 0.29 points, P<.001, Cohen d=0.63) groups. No interaction or main effects were seen in other self-care behaviors scores such as blood-glucose testing and medication adherence (Tables 2 and 3).

Table . Intention-to-treat (ITT) analysis: comparison of knowledge and behavior of diabetic foot self-care between the groups before and after the intervention (n=77; intervention group n=37 and control group n=40).

Group	Baseline, estimated mean (SE)	12 months, estimated mean (SE)	Within-group comparison, P value (Cohen d)	Time effect, P value (η^2)	Group effect, P value (η^2)	Interaction, P value (η^2)
Knowledge of diabetic foot care score						
Intervention	6.51 (0.31)	6.10 (0.29)	.2 (–0.38)	.01[a] (0.09)	.42 (0.01)	.39 (0.01)
Control	6.42 (0.27)	5.54 (0.36)	.02[a] (–0.21)	—[b]	—	—
Foot care and diabetes self-care behaviors						
Foot-care score						
Intervention	3.49 (0.19)	4.64 (0.36)	.001[a] (0.60)	<.001[a] (0.29)	.12 (0.04)	.98 (<0.001)
Control	3.04 (0.17)	4.16 (0.36)	.001[a] (0.62)	—	—	—
Blood-glucose testing score						
Intervention	4.77 (0.43)	5.16 (0.48)	.45 (–0.07)	.06 (0.05)	.07 (0.001)	.41 (0.009)
Control	4.30 (0.42)	5.34 (0.42)	.08 (0.03)	—	—	—
Medication adherence score						
Intervention	5.50 (0.35)	5.63 (0.42)	.83 (0.29)	.92 (0.0001)	.83 (0.0006)	.67 (0.002)
Control	5.60 (0.38)	5.38 (0.38)	.68 (0.12)	—	—	—

---

[a] P<.05.

[b] Em dashes indicate values that are identical for both groups because these results represent overall ANOVA model effects rather than group-specific estimates.

Table . Per-protocol (PP) analysis: comparison of knowledge and behavior of diabetic foot self-care between the groups before and after the intervention (n=66; intervention group n=34 and control group n=32).

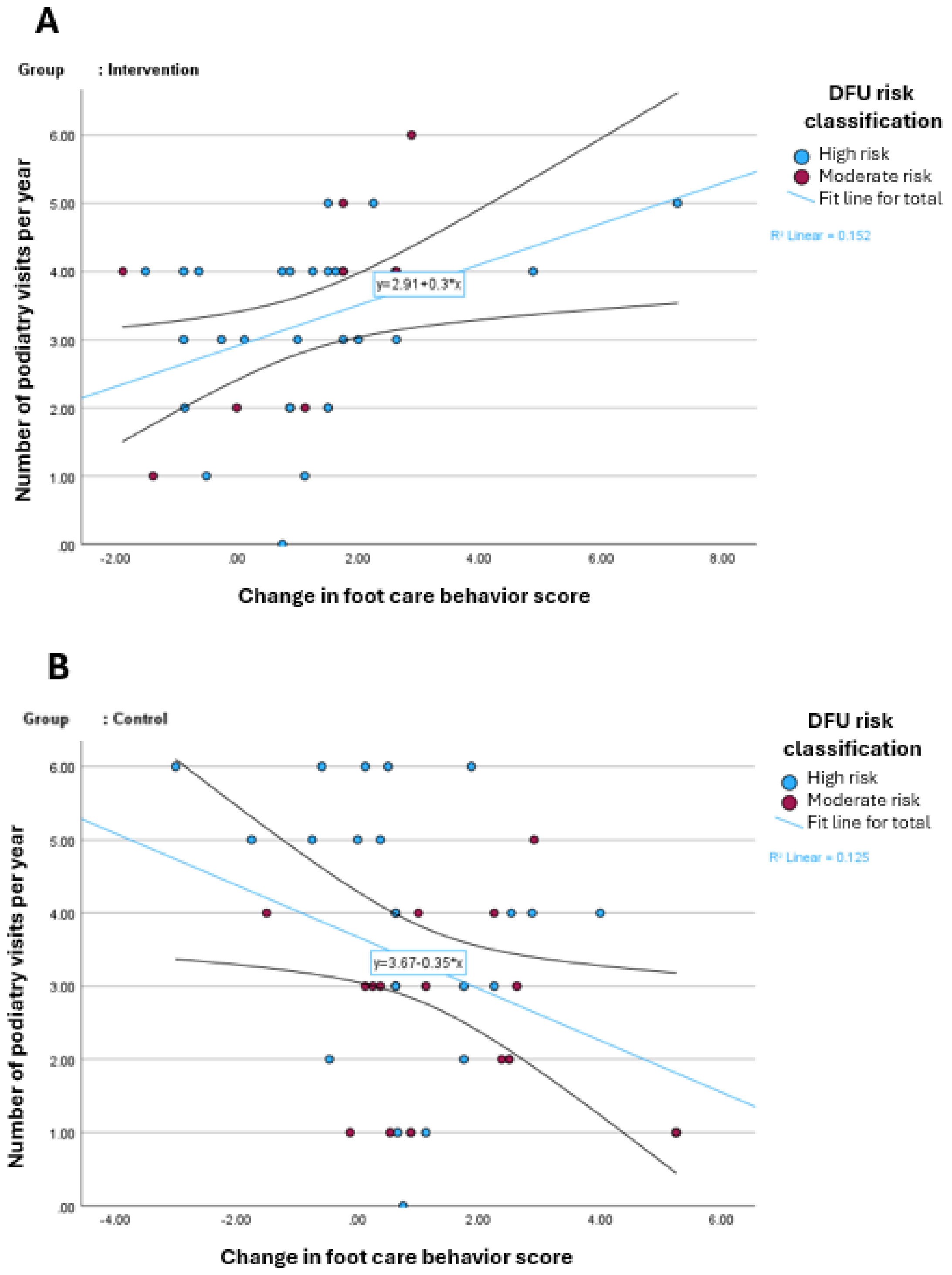
Group	Baseline, estimated mean (SE)	12 months, estimated mean (SE)	Within-group comparison, *P* value (Cohen *d*)	Time effect, *P* value (η²)	Group effect, *P* value (η²)	Interaction, *P* value (η²)
Knowledge of diabetic foot care score						
Intervention	6.62 (1.87)	6.15 (1.74)	.18 (–0.48)	.004[a] (0.12)	.49 (0.007)	.29 (0.017)
Control	6.63 (1.75)	5.63 (1.87)	.007[a] (–0.24)	—[b]	—	—
Foot care and diabetes self-care behaviors						
Foot-care score						
Intervention	3.51 (1.21)	4.63 (2.2)	<.001[a] (0.63)	<.001[a] (0.31)	.13 (0.034)	.97 (<0.001)
Control	3.03 (0.98)	4.14 (1.55)	<.001[a] (0.61)	—	—	—
Blood-glucose testing score						
Intervention	4.66 (2.71)	5.07 (2.91)	.41 (–0.07)	.09 (0.043)	.91 (0.0002)	.49 (0.01)
Control	4.3 (2.75)	5.29 (2.46)	.13 (–0.07)	—	—	—
Medication adherence score						
Intervention	5.35 (2.32)	5.61 (2.51)	.66 (–0.26)	.98 (<0.001)	.68 (0.003)	.79 (0.001)
Control	5.46 (2.54)	5.34 (2.26)	.71 (–0.15)	—	—	—

[a] *P*<.05.

[b] Em dashes indicate values that are identical for both groups because these results represent overall ANOVA model effects rather than group-specific estimates.

The mean of the number of podiatry follow-up visits over the 12 months was comparable between the ITT group (control: mean 3.27, SD 1.7 vs intervention: mean 3.24, SD 1.3 visits; *P*=.93) and PP group (control: mean 3.3, SD 1.6 vs intervention: mean 3.26, SD 1.3 visits; *P*=.83). Similarly, among participants at a high DFU risk, those in the intervention attended mean 3.52 (SD 1.2; ITT) and mean 3.64 (SD 1.1; PP) visits, while those in the control group attended mean 3.72 (SD 1.6; ITT) and mean 3.77 (SD 1.6; PP) visits. Among participants at a moderate DFU risk, the intervention group attended a mean of 2.25 (SD 1.48; ITT) visits, in both ITT and PP analyses, while the control group attended a mean of 2.53 (SD 1.6; ITT) visits and a mean of 2.76 (SD 1.58; PP) visits. The number of podiatry visits, however, was positively associated with the change in foot care behavior in the intervention group (ITT: *r*=0.390, *P*=.02; PP: *r*=0.408, *P*=.02), while the control group showed a negative correlation (ITT: *r*=−0.354, *P*=.03; PP: *r*=−0.402, *P*=.02; Figure 4). In ITT analysis, a significant negative association was observed in the control group between the number of podiatry visits and the change in the knowledge scores (*r*=−0.357; *P*=.02), but no such association was found in the intervention group. In the PP analysis, no significant associations were observed in either group. The frequency of new DFU during the course of the study was comparable between the 2 groups (ITT: control=17.5% vs intervention=21.6%, *P*=.64; PP: control=18.8% vs intervention=17.6%, *P*=.91). Both ITT and PP analyses revealed no statistically significant difference in the HbA_{1c} at 12 months between the intervention (ITT: mean 7.86%, SD 0.38%; PP: mean 7.93%, SD 1.08%) and the control (ITT: mean 8.28%, SD 0.36%; PP: mean 8.45%, SD 1.4%) groups, with a mean difference of 0.58% (95% CI −0.93 to 2.09; *P*=.44) in ITT, and 0.52% (95% CI −0.14 to 1.19; *P*=.12) in the PP analyses.

Figure 4. Association between the number of podiatry visits and the change in diabetic foot care behavior in the intervention group (**A**) and control group (**B**). DFU: diabetic foot ulceration.



Discussion

Principal Findings

This pilot study investigated the effectiveness of a multimedia educational approach in improving foot care knowledge and behavior among individuals with diabetes over a 12-month period. Overall, this study found that there was no long-term benefit to using multimedia educational tools in addition to standard usual podiatry care to improve participants' knowledge and behavior. However, post hoc comparisons showed that while both usual care and multimedia educational approaches enhanced the foot care behaviors of participants, the usual care settings resulted in a decline in their foot care knowledge, whereas the multimedia educational settings maintained it. The effect on glycemic control, however, was marginal and not clinically relevant (nonsignificant reduction in HbA_{1c} of only 0.52%). The consistency of results between the ITT and PP analyses supports the robustness of our findings and indicates that participant dropout did not substantially bias the outcomes.

This consistency increases confidence that the observed effects are likely to be representative of real-world settings.

Delivering the appropriate foot self-care education is a crucial strategy for reducing the risk of DFUs in people with diabetes [36]. Though there is currently no consensus regarding the optimal education method on diabetic foot care among health care professionals [37]. In light of technological advancements, several interventional studies using smartphone apps and media aids were conducted to educate patients about foot self-care [38]. But their effectiveness remains unclear. To the best of our knowledge, this is the first study that examined the long-term (12 mo) effects of multimedia educational tools on diabetic foot care including both audio-visual presentation and smartphone app. In support of our findings, a short-term (3 mo) RCT on 120 individuals with diabetes found that an audio-visual educational presentation, supplementary to standard diabetic care and regular reinforcement of education in outpatient clinic, significantly enhanced knowledge scores by +1.17 (95% CI 0.7 - 1.64, *P* value<.001) compared to those receiving standard care alone (+0.1, 95% CI −0.3 to 0.5, *P*=.62), whereas the practice scores were significantly improved in both groups (intervention: +1.6, 95% CI 1.09 - 2.11, *P*<.001; control: +0.48, 95% CI 0.16 - 0.8, *P*=.004) [39]. A 3-week educational program involving presentation slides, booklets, and group discussions significantly improved knowledge (intervention: mean 8.08, SD 0.88 vs control: mean 7.17, SD 1.91; *P*=.02) and behavior scores (intervention: mean 19.32, SD 6.48 vs control: mean 18.96, SD 7.36; *P*=.0001) compared to usual care [40].

Contrary to our findings regarding the use of a smartphone app, a recent RCT study involving 88 participants found that m-DAKBAS, an educational mobile diabetic foot care app that involved information, prevention, and management interfaces, significantly improved knowledge scores over a 6-month period (intervention: mean 16.83, SD 1.56 vs control: mean 15.05, SD 2.17; *P*=.0001), with no significant differences in behavior scores (intervention: mean 62.59, SD 7.76 vs control: mean 59.45, SD 10.53*;* *P*=.23) [41]. Furthermore, a 2-month use of the MobileDiabetes self-care app, which allowed patients to improve their self-care practices with flexibility in timing, location, and choices, has been shown to increase both patients' self-care knowledge and behavior by 17% and 22%, respectively [42]. In a 1-month RCT, an animation-supported app (Mobile Diabetic Foot Care Education; M-DFCE) that delivered 2 push alerts per week and included a cartoon animation providing basic education on daily foot care to prevent foot wounds was evaluated [43]. The experimental group outperformed the control group in terms of foot care behavior (mean difference: 11.28, SD 10.47 for intervention vs 0.6, SD 24.85 for control; *P* =.01) and knowledge (mean difference: 0.87, SD 1.21 for intervention vs 0.01, SD 1.25 for control; *P*=.002) when compared to baseline [43]. According to a recent systematic review, for patients with diabetes, using a mobile health app enhanced their awareness of the disease and their capacity for self-care in studies of durations of less than 6 months [44]. A study involving 58 patients with uncontrolled diabetes found that integrating self-management, through peer-supporting video, a quiz game, and a feedback system, with the Diabetes Care App for 5 weeks, improved foot care behavior significantly (*P*<.01) when compared to usual care [45]. Similar improvements in foot care practices were observed on patients with type 2 diabetes after 1 month [46] and 3 months [47] of using self-management integration with smartphone apps. Furthermore, the use of smartphone apps and alerts interventions has been associated with improved glycemic control [48], which is contrary to our findings. We found a minor and comparable decrease in HbA_{1c} among the 2 groups, which might be reflected by their scores of the blood glucose-testing and medication adherence. While multimedia education on foot self-care is essential for improving the knowledge and behavior of individuals at a moderate-to-high risk of DFUs, it did not have a significant impact on glycemic control or long-term behavior changes.

In this study, there was no discernible difference in the overall incidence of DFU across the groups, although the overall incidence and thus the study power were low. This may suggest that the use of multimedia aids in foot self-care does not offer additional benefits above the usual care education. In agreement with our findings, a systematic review of 6 RCTs found insufficient evidence for the benefit of an integrated care approach, which involved combining multiple DFU prevention strategies at different levels of care (including the patient, health care provider, and health care structure), indicating the need for more high-quality studies [22]. Another systematic review of 11 RCTs, on the other hand, showed that educational technologies were protective against the incidence of lower limb amputations (relative risk=0.53, 95% CI 0.31 - 0.90; *P*=.02) and DFU (relative risk=0.40, 95% CI 0.18 - 0.90; *P*=.03), despite the lack of evidence of certainty assessment [49]. Other studies suggested that a more focused and intensive educational approach should be adopted to reduce the incidence of DFU. An RCT on people with type 2 diabetes found that 2-hour focused education sessions, including practical exercises on foot care behaviors, were effective in preventing the incidence of DFU during 6-month follow-ups compared to the control group (incidence of DFU=0% vs 10%; *P*=.01) [50]. Similarly, a quasi-experimental study found that combining educational sessions with foot assessment and care reduced the recurrence rate of DFU to 13.3% compared to 33.3% in the control group [51]. Intensive education approaches that included training and customized footwear have been shown to reduce the incidence rate of new DFU compared to usual care by 18% and 31%, respectively [52]. The results from a systematic review and meta-analysis of RCTs demonstrated that, in comparison to the control group, an intensive educational approach—a 45- to 1-hour education session reinforced by written instructions—significantly decreased the risk of DFU incidence (relative risk 0.37, 95% CI 0.14-1.01; *P*=.05) [53]. Nonetheless, the study showed that there was a substantial degree of variability in the studies, with 91% heterogeneity [53]. Thus, multimedia aids may not offer additional benefits beyond usual care in reducing the incidence of DFU among individuals at a moderate-high risk of DFUs, suggesting the need for more targeted educational approaches. Although our study did not find a significant effect on self-care behavior, this should not be interpreted as evidence against the long-term effectiveness of all media-based tools, which may require extended follow-ups, greater interactivity, or more tailored content to achieve meaningful and sustained behavioral change. In addition, future studies should consider evaluating the effectiveness of media-based educational tools in lower-risk populations, where such interventions may have a greater impact on preventive behaviors and outcomes.

Strengths and Limitations

The strength of this pilot study is the long-term duration of the intervention at 12 months. However, a major limitation of this study is the unexpected COVID-19 global pandemic and associated government restrictions, which significantly impacted study conduct and led to lower enrollment and higher dropout rates than expected. This was potentially because, as a high-risk group for COVID-19, participants were reluctant to increase their contact with others unnecessarily. Although there was a higher dropout among participants with a high risk, particularly in the control group, our ITT analysis yielded results consistent with the PP analysis, suggesting that differential attrition did not substantially bias the study findings. The small sample size limits the generalizability of our findings, and thus, future studies with larger and more diverse populations are needed to validate these results and better evaluate the effectiveness of app-based interventions. Another limitation is that the MyU app lacked interactive features, which may have reduced its appeal—particularly among older participants [54,55]—although this was not formally assessed through usability testing or structured feedback during the study.

Conclusions and Future Directions

In conclusion, the study findings revealed that a multimedia education approach alone to improve foot self-care knowledge and behavior was unfeasible and not engaging for long-term use in people at a moderate-high DFU risk. However, the reinforcement of educational material in the follow-up podiatry visit might be effective in achieving persistent changes in foot care behavior of this patient group. Given the scope of this study, further confirmative RCT studies, with a larger sample size, are needed.

Acknowledgments

The authors would like to thank the participants for volunteering to take part in the study. The authors would also like to acknowledge the support of the nurses, clinical laboratory, and podiatry department at the Dasman Diabetes Institute.

Funding

The study has been funded by the Dasman Diabetes Institute, Kuwait.

Data Availability

The data supporting the results of this study are available from the corresponding author upon reasonable request, with a data transfer agreement if required.

Authors' Contributions

Conceptualization: EAO, GM.
Data curation: AC, GM, KSK, MA.
Data validation: AA, SRG.
Formal analysis: AA.
Funding acquisition: EAO, GM.
Project administration: AC, EAO, KSK, MA, SRG.
Resources: AC, GM, KSK, MA.
Study design: EAO, GM, MA.
Supervision: AC, EAO, KSK, MA, SRG.
Visualization: AA, SRG.
Writing – original draft: AA.
Writing – review & editing: AA, AC, EAO, GM, KSK, MA, SRG.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH checklist.
[PDF File, 2620 KB - humanfactors_v13i1e78261_app1.pdf]

References

1. Ong KL, Stafford LK, McLaughlin SA. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2023 Jul 15;402(10397):203-234. [doi: 10.1016/S0140-6736(23)01301-6] [Medline: 37356446]

2. Magliano DJ, Boyko EJ. IDF DIABETES ATLAS, 10th edition: International Diabetes Federation; 2021. URL: http://www.ncbi.nlm.nih.gov/books/NBK581934 [accessed 2025-12-13]
3. State of the nation (England 2016): time to take control of diabetes. Diabetes UK. 2016. URL: https://www.diabetes.org.uk/node/1753 [accessed 2025-12-13]
4. Reiber GE, Vileikyte L, Boyko EJ, et al. Causal pathways for incident lower-extremity ulcers in patients with diabetes from two settings. Diabetes Care 1999 Jan;22(1):157-162. [doi: 10.2337/diacare.22.1.157] [Medline: 10333919]
5. Abbott CA, Garrow AP, Carrington AL, et al. Foot ulcer risk is lower in South-Asian and African-Caribbean compared with European diabetic patients in the U.K.: the North-West diabetes foot care study. Diabetes Care 2005 Aug;28(8):1869-1875. [doi: 10.2337/diacare.28.8.1869] [Medline: 16043725]
6. Singh N, Armstrong DG, Lipsky BA. Preventing foot ulcers in patients with diabetes. JAMA 2005 Jan 12;293(2):217-228. [doi: 10.1001/jama.293.2.217] [Medline: 15644549]
7. Boulton AJM. The diabetic foot: grand overview, epidemiology and pathogenesis. Diabetes Metab Res Rev 2008;24 Suppl 1:S3-S6. [doi: 10.1002/dmrr.833] [Medline: 18442166]
8. Pecoraro RE, Reiber GE, Burgess EM. Pathways to diabetic limb amputation. Basis for prevention. Diabetes Care 1990 May;13(5):513-521. [doi: 10.2337/diacare.13.5.513] [Medline: 2351029]
9. Siitonen OI, Niskanen LK, Laakso M, Siitonen JT, Pyörälä K. Lower-extremity amputations in diabetic and nondiabetic patients. A population-based study in eastern Finland. Diabetes Care 1993 Jan;16(1):16-20. [doi: 10.2337/diacare.16.1.16] [Medline: 8422771]
10. Trautner C, Haastert B, Giani G, Berger M. Incidence of lower limb amputations and diabetes. Diabetes Care 1996 Sep;19(9):1006-1009. [doi: 10.2337/diacare.19.9.1006] [Medline: 8875099]
11. Armstrong DG, Swerdlow MA, Armstrong AA, Conte MS, Padula WV, Bus SA. Five year mortality and direct costs of care for people with diabetic foot complications are comparable to cancer. J Foot Ankle Res 2020 Mar 24;13(1):16. [doi: 10.1186/s13047-020-00383-2] [Medline: 32209136]
12. Lim TS, Finlayson A, Thorpe JM, et al. Outcomes of a contemporary amputation series. ANZ J Surg 2006 May;76(5):300-305. [doi: 10.1111/j.1445-2197.2006.03715.x] [Medline: 16768686]
13. Schaper NC, van Netten JJ, Apelqvist J, et al. Practical guidelines on the prevention and management of diabetic foot disease (IWGDF 2019 update). Diabetes Metab Res Rev 2020 Mar;36 Suppl 1(S1):e3266. [doi: 10.1002/dmrr.3266] [Medline: 32176447]
14. Ezzatvar Y, García-Hermoso A. Global estimates of diabetes-related amputations incidence in 2010-2020: a systematic review and meta-analysis. Diabetes Res Clin Pract 2023 Jan;195:110194. [doi: 10.1016/j.diabres.2022.110194] [Medline: 36464091]
15. Raju BN, Mateti UV, Mohan R, et al. Educational interventions and its impact on the treatment outcomes of diabetic foot ulcer patients. J Diabetol 2022;13(1):1-7. [doi: 10.4103/jod.jod_99_21]
16. Yıldırım Ayaz E, Dincer B, Oğuz A. The effect of foot care education for patients with diabetes on knowledge, self-efficacy and behavior: systematic review and meta-analysis. Int J Low Extrem Wounds 2022 Sep;21(3):234-253. [doi: 10.1177/15347346221109047] [Medline: 35711163]
17. Saltar L, Sahar J. The intervention of foot care education in the prevention of diabetic foot ulcers: a literature review. 2020 Presented at: International Summit on Science, Technology, and Humanity (ISETH); Dec 15-16, 2020 URL: https://proceedings.ums.ac.id/iseth/article/view/1276 [accessed 2021-08-25]
18. Sampson EO, Abdul Manaf R, Ismail S, Shahar HK, Udeani TK. Improving foot self-care practices through health education intervention programs among diabetic patients: a systematic review. Malays J Med Health Sci 2023;19(4):315-325. [doi: 10.47836/mjmhs.19.4.44]
19. Dorresteijn JAN, Kriegsman DMW, Assendelft WJJ, Valk GD. Patient education for preventing diabetic foot ulceration. Cochrane Database Syst Rev 2014 Dec 16;2014(12):CD001488. [doi: 10.1002/14651858.CD001488.pub5] [Medline: 25514250]
20. Campbell FA, Goldman BD, Boccia ML, Skinner M. The effect of format modifications and reading comprehension on recall of informed consent information by low-income parents: a comparison of print, video, and computer-based presentations. Patient Educ Couns 2004 May;53(2):205-216. [doi: 10.1016/S0738-3991(03)00162-9] [Medline: 15140461]
21. Murphy PW, Chesson AL, Walker L, Arnold CL, Chesson LM. Comparing the effectiveness of video and written material for improving knowledge among sleep disorders clinic patients with limited literacy skills. South Med J 2000 Mar;93(3):297-304. [Medline: 10728518]
22. Hoogeveen RC, Dorresteijn JAN, Kriegsman DMW, Valk GD. Complex interventions for preventing diabetic foot ulceration. Cochrane Database Syst Rev 2015 Aug 24;2015(8):CD007610. [doi: 10.1002/14651858.CD007610.pub3] [Medline: 26299991]
23. Pratiwi LD, Haryanto J, Wahyudi AS. The effect of foot self care and diabetes self-management mobile application in preventing foot ulcer recurrence: a systematic review study. Malays J Med Health Sci 2023;19:109-116 [FREE Full text]
24. Obilor HN, Achore M, Woo K. Use of information communication technology tools in diabetic foot ulcer prevention programs: a scoping review. Can J Diabetes 2022 Jul;46(5):535-548. [doi: 10.1016/j.jcjd.2021.11.009] [Medline: 35739048]

25. Bus SA, van Netten JJ. A shift in priority in diabetic foot care and research: 75% of foot ulcers are preventable. Diabetes Metab Res Rev 2016 Jan;32 Suppl 1(S1):195-200. [doi: 10.1002/dmrr.2738] [Medline: 26452160]
26. Process of translation and adaptation of instruments. World Health Organization. 2016. URL: https://www.scirp.org/reference/referencespapers?referenceid=3514337 [accessed 2025-07-06]
27. Diabetic foot problems: prevention and management. National Institute for Health and Care Excellence. 2015 Aug 26. URL: https://www.nice.org.uk/guidance/ng19 [accessed 2025-07-03]
28. Management of diabetes: a national clinical guideline. Scottish Intercollegiate Guidelines Network. URL: https://www.sign.ac.uk/assets/sign116.pdf [accessed 2025-12-09]
29. American Diabetes Association. Microvascular complications and foot care: standards of medical care in diabetes—2018. Diabetes Care 2018 Jan;41(Suppl 1):S105-S118. [doi: 10.2337/dc18-S010] [Medline: 29222381]
30. Sheridan SL, Viera AJ, Krantz MJ, et al. The effect of giving global coronary risk information to adults: a systematic review. Arch Intern Med 2010 Feb 8;170(3):230-239. [doi: 10.1001/archinternmed.2009.516] [Medline: 20142567]
31. Pollock RD, Unwin NC, Connolly V. Knowledge and practice of foot care in people with diabetes. Diabetes Res Clin Pract 2004 May;64(2):117-122. [doi: 10.1016/j.diabres.2003.10.014] [Medline: 15063604]
32. Rheeder P, Venn M, de Korte E, van Zyl D. Knowledge of foot care in people with diabetes in a tertiary care setting. J Endocrinol Metab Diabetes S Afr 2008 Dec;13(3):105-108. [doi: 10.1080/22201009.2008.10872180]
33. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950. [doi: 10.2337/diacare.23.7.943] [Medline: 10895844]
34. AlJohani KA, Kendall GE, Snider PD. Psychometric evaluation of the Summary of Diabetes Self-Care Activities-Arabic (SDSCA-Arabic): translation and analysis process. J Transcult Nurs 2016 Jan;27(1):65-72. [doi: 10.1177/1043659614526255] [Medline: 24841470]
35. Baba M, Duff J, Foley L, Davis WA, Davis TME. A comparison of two methods of foot health education: the Fremantle Diabetes Study Phase II. Prim Care Diabetes 2015 Apr;9(2):155-162. [doi: 10.1016/j.pcd.2014.05.004] [Medline: 24929632]
36. Schaper NC, van Netten JJ, Apelqvist J, et al. Practical guidelines on the prevention and management of diabetic foot disease (IWGDF 2019 update). Diabetes Metab Res Rev 2020 Mar;36(S1):e3266. [doi: 10.1002/dmrr.3266] [Medline: 32176447]
37. Jeffcoate WJ, Vileikyte L, Boyko EJ, Armstrong DG, Boulton AJM. Current challenges and opportunities in the prevention and management of diabetic foot ulcers. Diabetes Care 2018 Apr;41(4):645-652. [doi: 10.2337/dc17-1836] [Medline: 29559450]
38. Kustini K, Sari Y. Use of mobile applications in increasing knowledge of diabetes mellitus foot care. J Med Health Stud 2023;4(6):21-24. [doi: 10.32996/jmhs.2023.4.6.3]
39. Rahaman HS, Jyotsna VP, Sreenivas V, Krishnan A, Tandon N. Effectiveness of a patient education module on diabetic foot care in outpatient setting: an open-label randomized controlled study. Indian J Endocrinol Metab 2018;22(1):74-78. [doi: 10.4103/ijem.IJEM_148_17] [Medline: 29535941]
40. Beiranvand S, Fayazi S, Asadizaker M. Effect of educational programs on the knowledge, attitude, and practice of foot care in patients with diabetes. Jundishapur J Chronic Dis Care 2015;4(2). [doi: 10.5812/jjcdc.26540]
41. Kilic M, Karadağ A. Developing and evaluating a mobile foot care application for persons with diabetes mellitus: a randomized pilot study. Wound Manag Prev 2020 Oct;66(10):29-40. [doi: 10.25270/wmp.2020.10.2940] [Medline: 33048829]
42. Guo SHM, Chang HK, Lin CY. Impact of mobile diabetes self-care system on patients' knowledge, behavior and efficacy. Comput Ind 2015 May;69:22-29. [doi: 10.1016/j.compind.2014.11.001]
43. Dincer B, Bahçecik N. The effect of a mobile application on the foot care of individuals with type 2 diabetes: a randomised controlled study. Health Educ J 2021 Jun;80(4):425-437. [doi: 10.1177/0017896920981617]
44. Sadler S, Gerrard J, Searle A, et al. The use of mHealth apps for the assessment and management of diabetes-related foot health outcomes: systematic review. J Med Internet Res 2023 Oct 4;25:e47608. [doi: 10.2196/47608] [Medline: 37792467]
45. Firdaus M, Jittanoon P, Boonyasopun U, Che Hasan MK. The effect of mHealth program on behavior modification and health outcomes among patients with diabetes: a randomized controlled trial study. Belitung Nurs J 2023;9(5):437-447. [doi: 10.33546/bnj.2664] [Medline: 37901368]
46. Chen SM, Jo Wu CJ. Investigating the effects of digital foot self-management program on enhancing self-efficacy and self-care behavior among community-dwelling older adults with type 2 diabetes: a randomized controlled trial. Digit Health 2023;9:20552076231220791. [doi: 10.1177/20552076231220791] [Medline: 38107978]
47. Suniyadewi NW, Nursalam N, Sufyanti Arief Y, et al. Effect of android-based mobile diabetic foot early self-assessment on diabetic foot prevention behaviors of Indonesian patients with type 2 diabetes. J Holist Nurs Midwifery 2024;34(4):309-315. [doi: 10.32598/jhnm.34.4.2633]
48. Moschonis G, Siopis G, Jung J, et al. Effectiveness, reach, uptake, and feasibility of digital health interventions for adults with type 2 diabetes: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health 2023 Mar;5(3):e125-e143. [doi: 10.1016/S2589-7500(22)00233-3] [Medline: 36828606]

49. Lira JAC, Rocha Á, Bezerra SMG, Nogueira PC, Santos A, Nogueira LT. Effects of educational technologies on the prevention and treatment of diabetic ulcers: a systematic review and meta-analysis. Rev Lat Am Enfermagem 2023;31:e3944. [doi: 10.1590/1518-8345.6628.3944] [Medline: 37341258]
50. Monami M, Zannoni S, Gaias M, Nreu B, Marchionni N, Mannucci E. Effects of a short educational program for the prevention of foot ulcers in high-risk patients: a randomized controlled trial. Int J Endocrinol 2015;2015(1):615680. [doi: 10.1155/2015/615680] [Medline: 26448748]
51. Pratama K, Fahrain J, et al. Prevention strategy for ulcer recurrence in patients with type II diabetes mellitus: a quasi-experimental study. Iran J Nurs Midwifery Res 2023;28(2):139-143. [doi: 10.4103/ijnmr.ijnmr_432_21] [Medline: 37332375]
52. Bharat Kotru SK. Intervention of diabetes foot care practices on the prevention of new diabetic foot ulcers in patients with type 2 diabetes mellitus. J Diabetes Metab 2025;06(2):1-6. [doi: 10.4172/2155-6156.1000494]
53. Adiewere P, Gillis RB, Imran Jiwani S, Meal A, Shaw I, Adams GG. A systematic review and meta-analysis of patient education in preventing and reducing the incidence or recurrence of adult diabetes foot ulcers (DFU). Heliyon 2018 May;4(5):e00614. [doi: 10.1016/j.heliyon.2018.e00614] [Medline: 29872752]
54. Steinert A, Eicher C, Haesner M, Steinhagen-Thiessen E. Effects of a long-term smartphone-based self-monitoring intervention in patients with lipid metabolism disorders. Assist Technol 2020;32(2):109-116. [doi: 10.1080/10400435.2018.1493710] [Medline: 29944463]
55. Su J, Dugas M, Guo X, Gao GG. Influence of personality on mHealth use in patients with diabetes: prospective pilot study. JMIR mHealth mHealth 2020 Aug 10;8(8):e17709. [doi: 10.2196/17709] [Medline: 32773382]

Abbreviations

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
DFU: diabetic foot ulceration
HbA_{1c}: hemoglobin A_{1c}
ITT: intention-to-treat
NICE: National Institute for Health and Care Excellence
PP: per-protocol
RCT: randomized controlled trial
SDSCA: Summary of Diabetes Self-Care Activities

Edited by A Kushniruk; submitted 29.May.2025; peer-reviewed by C Law, DV Coppini, Y Gao; revised version received 15.Sep.2025; accepted 14.Oct.2025; published 13.Jan.2026.

Please cite as:
Alawadhi A, Khan KS, Messenger G, Cusworth A, Assi M, Gray SR, Al-Ozairi E
Long-Term Effects of Multimedia Education and “Foot Alerts” Through the MyU App on Diabetic Foot Care in Individuals With Moderate-High Risk: Pilot Randomized Controlled Trial
JMIR Hum Factors 2026;13:e78261
URL: https://humanfactors.jmir.org/2026/1/e78261
doi:10.2196/78261

Utilization of AI Among Medical Students and Development of AI Education Platforms in Medical Institutions: Cross-Sectional Study

Xiaokang Shi[1,2], BS; Zewu Jiang[1], MM; Li Xiong[1], BS; Ka-Chun Siu[3], PhD; Zhen Chen[1,2], MD, PhD

[1]Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Building 9, 3rd Floor, No. 160, Pujian Road, Pudong New Area, Shanghai, China

[2]School of Medicine, Tongji University, Shanghai, China

[3]Department of Health and Rehabilitation Sciences, College of Allied Health Professions, University of Nebraska Medical Center, Omaha, NE, United States

Corresponding Author:
Zhen Chen, MD, PhD
Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Building 9, 3rd Floor, No. 160, Pujian Road, Pudong New Area, Shanghai, China

Abstract

Background: The emergence of artificial intelligence (AI) is driving digital transformation and reshaping medical education in China. Numerous medical schools and institutions are actively implementing AI tools for case-based learning, literature analysis, and lecture support. This expanding application is accelerating the adoption of localized AI platforms, which are poised to become integral components in the coming years.

Objective: The primary aim of this study was to investigate the current use of AI tools among medical students, including usage frequency, commonly used platforms, and purposes of use. The second aim was to explore students' needs and expectations toward AI-powered medical education platforms by collecting and assessing student feedback, and to identify practical requirements across disciplines and academic stages to inform more effective platform design.

Methods: Based on the task-technology fit model and 5 hypotheses, an anonymous online questionnaire was conducted to assess AI usage in learning, gather student feedback on AI-powered medical education platforms, and evaluate expected functionalities. The survey was conducted from March 1 to May 31, 2025, using a convenience sampling method to recruit medical students from various disciplines across Shanghai, China. The sample size was determined at 422, accounting for a 10% rate of invalid responses. The questionnaire was developed and distributed online via Wenjuanxing and promoted through WeChat groups and in-person interviews. Data analysis was conducted employing IBM SPSS Statistics (v 27.0).

Results: A total of 428 valid questionnaires were collected. The average frequency of AI-assisted learning among medical students was 5.06 (SD 2.05) times per week. Over 90% (388/428) of the students used more than 2 AI tools in their daily tasks. Students from different disciplines, educational stages, and academic systems demonstrated different usage patterns and expectations for AI-powered medical education platforms.

Conclusions: AI technology is widely accepted by medical students and is extensively applied across various aspects of medical education. Significant differences are observed in usage patterns across disciplines, educational stages, and academic systems. Understanding the actual needs of students is crucial for the construction of AI-powered medical education platforms.

(JMIR Hum Factors 2026;13:e81652) doi:10.2196/81652

KEYWORDS

AI chatbots; artificial intelligence; cross-sectional study; medical education; medical schools; medical students; technology acceptance

Introduction

The rapid development of artificial intelligence (AI) has profoundly accelerated the digital transformation of medical education worldwide. AI demonstrates significant potential across multiple domains of medical education, including case-based teaching, literature analysis, and lecture support [1,2].

Globally, previous studies have documented the successful integration of large language models (LLMs) and conversational agents in medical education, significantly enhancing teaching effectiveness [3,4]. Gilson et al [5] reported that ChatGPT and other LLMs have been deployed to simulate clinical reasoning sessions, automate feedback on student essays, and generate customized practice questions, demonstrating measurable gains in diagnostic accuracy and pedagogical efficiency. Simultaneously, students' mastery of both theoretical knowledge and practical skills has markedly improved, resulting in better learning outcomes and clinical performance. These advances address existing challenges in medical education and open up promising pathways for its future development [6,7].

In China, supportive national policies and rapid technological advances have jointly facilitated the localized application of AI in medical education [8]. The national “AI Plus” Implementation Guidelines outline the strategic direction for the deep integration of AI technologies with public welfare services [9]. Meanwhile, Shanghai’s pioneering Medical Artificial Intelligence Work Plan specifically proposes establishing an “intelligent medical education and training” platform, emphasizing the development of smart training environments using generative AI technologies [10]. This provides clear policy support and practical guidance for higher education institutions to develop localized AI−powered medical education platforms [11,12].

With the widespread adoption of AI, students’ learning methods, habits, and institutional teaching models have been rapidly reshaped. However, systematic data on medical students’ current AI usage patterns and their practical needs across different educational stages and disciplines remain limited. A deeper understanding of these aspects is crucial for effectively guiding the development and optimization of future AI−powered medical education platforms. Therefore, this study aimed to collect medical students’ current use of AI and their practical needs across different educational stages and academic disciplines. We also explored effective strategies for developing AI-powered medical education platforms, with the goal of providing recommendations to guide the development and optimization of future AI−powered medical education platforms.

Methods

Participants and Procedures

This study employed a convenience sampling method to conduct a cross-sectional survey among medical students from various medical universities in Shanghai, China, with data collected from March 1 to May 31, 2025.

The inclusion criteria were (1) current enrollment in a medical program and (2) provision of informed consent to participate. The exclusion criteria were (1) nonmedical students, (2) students not currently enrolled in any program, (3) students attending medical schools outside Shanghai, and (4) students who declined participation.

Theoretical Framework and Hypotheses

This study is based on the task-technology fit (TTF) model. Its core proposition is that technology’s effectiveness is determined not by its attributes or user attitudes alone but by the fit between technological functionalities and user task requirements.

Based on the TTF framework and a review of relevant literature [13,14], this study proposes 5 hypotheses. These hypotheses aim to examine differences in usage status, functional needs, and expectations regarding AI-powered medical education platforms among medical students across different disciplines, educational stages, and program types while identifying factors with a significant influence.

- H1: Perceived task-technology fit has a significant positive impact on platform satisfaction.
- H2: Disciplinary background moderates the relationship between TTF and platform satisfaction.
- H3: Program type (full-time vs part-time) moderates the relationship between TTF and platform satisfaction.
- H4: Educational stage moderates the relationship between TTF and platform satisfaction.
- H5: The frequency of AI usage has a significant positive impact on platform satisfaction.

Instrument Pretesting and Validation

Guided by these 5 hypotheses, we designed a concise online questionnaire. Prior to the formal survey, a pilot test was conducted through in-person interviews with 23 postgraduate clinical medicine students from the same institution to evaluate content validity, item clarity, and internal consistency reliability. The expected platform functions scale (Cronbach α=0.825, items with zero variance were excluded) and the most frequently used AI platforms scale (Cronbach α=0.858, items with zero variance were excluded) showed high internal consistency.

Questionnaire Design

Based on the feedback from the pilot study, we made appropriate revisions to certain items and their phrasing in the questionnaire. Our questionnaire covered 3 sections with 13 items (see the questionnaire in Multimedia Appendix 1): the first section is general information, including age, gender, major, school, educational stage, and academic program type; the second section is the current use of AI tools, including the frequency of use (average per week), preferred platforms, and usage purposes (eg, theoretical learning, literature assistance, among others); and the third section is current status and expectations for AI-powered medical education platforms (whether the institution has developed an AI-powered medical education platform, satisfaction with the platform, and expected functions of the future platform) [13].

To ensure data completeness, all questionnaire items were set as mandatory, and participants were required to complete all the questions before submission. The survey platform would automatically record device type and completion time.

Ethical Considerations

This study received an ethics exemption (EX-2025 - 017) from the Medical Ethics Committee of Renji Hospital, Shanghai Jiao Tong University School of Medicine, as it utilizes anonymized data, operates under standard informed consent protocols, and involves no sensitive biological materials or procedures.

We obtained informed consent from all student participants before the survey, providing full details about the study’s purpose, procedures, and privacy protections. No compensation was provided, as the study involved minimal burden and no anticipated harm. All the data were strictly protected to ensure confidentiality and prevent any risk of information leakage. To this end, access was restricted to authorized research team members, and the data were used solely for analysis and reporting within this study.

Sample Size

To ensure adequate statistical power and precision for the intended analyses, we used a standard sample size calculation formula [15]. Assuming a 95% CI, a margin of error of 0.05, and an expected population proportion of 0.5, the minimum required sample size was calculated to be 384. Drawing on previous studies [16,17] and to improve the generalizability of the results, we further accounted for a 10% invalid questionnaire rate, resulting in a final target sample size of 422.

$n=Z^2 \cdot P \cdot (1-P)E^2$ (95% CI, Z=1.96, E=0.05, P=.50)

Finite population correction was not applied due to the use of convenience sampling and structural heterogeneity across institutions, which precluded the definition of a single unified sampling frame.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics (version 27.0). Descriptive statistics were computed using appropriate measures for each variable type: continuous variables were summarized with means and SDs, while categorical variables were presented as frequencies and percentages.

For group comparisons involving categorical variables, chi-square tests were employed. Multiple response analyses were conducted using multiple response sets combined with chi-square tests, with Bonferroni correction applied to account for multiple comparisons. For ordinal data or continuous variables violating normality assumptions, non-parametric tests (Mann-Whitney *U* for 2-group comparisons and Kruskal-Wallis *H* for multigroup comparisons) were utilized, with post hoc analyses performed where appropriate.

ANOVA was used for comparing continuous variables across multiple groups, while MANOVA was employed for analyses involving multiple continuous dependent variables. Multivariable analyses included linear regression for continuous outcomes and logistic regression for binary outcomes.

All statistical tests used a 2-tailed significance threshold of $P<.05$, with appropriate corrections for multiple testing implemented where necessary [18].

Results

Participants' Characteristics

A total of 440 questionnaires were collected. After excluding 12 responses from nonmedical students, students at institutions outside Shanghai, and nonenrolled individuals, 428 valid questionnaires were retained, yielding an effective response rate of 97.3%. The questionnaires were returned with complete and valid responses, thus containing no missing data.

Participants were drawn from 7 medical schools in Shanghai, with 188 (43.92%) male participants and 240 (56.07%) female participants, and a median age of 22 (IQR 20.07-24.62) years. Among them, 223 (52.10%) were undergraduate students, 174 (40.65%) were master's students, and 31 (7.24%) were doctoral students. The sample covered 8 major disciplines: clinical medicine, basic medicine, rehabilitation therapy, nursing, public health and epidemiology, pharmacy, traditional Chinese medicine, and medical engineering.

Frequency of AI Use Among Medical Students

Most respondents reported regular use of AI tools in academic work and daily tasks, with a mean usage frequency of 5.06 (SD 2.05) times per week. The mean (SD) frequency among undergraduates, master's students, and doctoral students was 5.09 (1.97), 4.99 (2.11), and 5.19 (2.27) times per week, respectively, with no significant differences based on educational stage or gender (Table 1).

Table . Current usage of artificial intelligence (AI) in medical students' learning

Variables	Educational stage			Chi-square (*df*)	Cramer *V*	*P* value[a]	Gender		Chi-square (*df*)	Cramer *V*	*P* value[a]	Total (n=428), n (%)
	Undergraduate students (n=223), n (%)	Master's students (n=174), n (%)	Doctoral students (n=31), n (%)				Male (n=188), n (%)	Female (n=240), n (%)				
Average weekly use of AI for learning				7.97 (6)	0.10	.24			6.29 (3)	0.12	.10	
≤1	15 (6.73)	21 (12.07)	4 (12.90)				16 (8.51)	24 (10)				40 (9.35)
2 - 3	58 (26.01)	34 (19.54)	6 (19.35)				36 (19.15)	62 (25.83)				98 (22.90)
4 - 6	52 (23.32)	44 (25.29)	4 (12.90)				40 (21.28)	60 (25.00)				100 (23.36)
≥7	98 (43.95)	75 (43.10)	17 (54.84)				96 (51.06)	94 (39.17)				190 (44.39)
AI platform used by medical students												
DeepSeek	213 (95.52)	163 (93.68)	28 (90.32)	1.67 (2)	0.06	.43	176 (93.62)	228 (95.00)	0.38 (1)	0.03	.54	404 (94.39)
Doubao	138 (61.88)	99 (56.90)	13 (41.94)	4.74 (2)	0.11	.09	97 (51.60)	153 (63.75)	6.41 (1)	0.12	.01	250 (58.41)
Kimi	127 (56.95)	108 (62.07)	13 (41.94)	4.57 (2)	0.10	.10	98 (52.13)	150 (62.50)	4.66 (1)	0.10	.03	248 (57.94)
ChatGPT	156 (69.96)	103 (59.20)	21 (67.74)	5.08 (2)	0.11	.08	130 (69.15)	150 (62.50)	2.06 (1)	0.07	.15	280 (65.42)
Chat-GLM	28 (12.56)	14 (8.05)	0 (0)	5.88 (2)	0.12	.05	17 (9.04)	25 (10.42)	0.23 (1)	0.02	.64	42 (9.81)
Claude	24 (10.76)	9 (5.17)	3 (9.68)	4.03 (2)	0.10	.13	22 (11.70)	14 (5.83)	4.71 (1)	0.11	.03	36 (8.41)
Gemini	29 (13.00)	13 (7.47)	0 (0)	7.02 (2)	0.13	.03	27 (14.36)	15 (6.25)	7.84 (1)	0.14	.005	42 (9.81)
Other[b]	22 (9.87)	11 (6.32)	1 (3.23)	2.70 (2)	0.08	.26	16 (8.51)	18 (7.50)	0.15 (1)	0.02	.70	34 (7.94)
Number of AI platforms used by medical students				17.36 (14)	0.14	.24			3.86 (7)	0.10	.80	
2	53 (23.77)	45 (25.86)	8 (25.81)				46 (24.47)	60 (25.00)				106 (24.77)
≥2	210 (94.17)	153 (87.93)	25 (80.65)				166 (88.30)	222 (92.50)				388 (90.65)
≥3	157 (70.40)	108 (62.07)	17 (54.83)				120 (63.83)	162 (67.50)				282 (65.89)

[a] Chi-square test and Bonferroni correction were applied for multiple comparisons (α=.00625).

[b] AI platforms, including Qwen Chat, ERNIE Bot, Tencent Yuanbao, Poe, and Grok, were grouped together for analysis due to their relatively small sample sizes.

AI Platforms Used by Medical Students

In terms of AI platform selection, most medical students favor mainstream tools such as DeepSeek (n=404, 94.39%), Doubao (n=250, 58.41%), and ChatGPT (n=280, 65.42%) for their daily study and work. Survey results further indicated that the current use of multiple AI platforms has become a common practice among medical students, with over 90% (388/428) of the students reporting the use of 2 or more platforms, and over 60% (282/428) of the students reported using 3 or more platforms (Table 1).

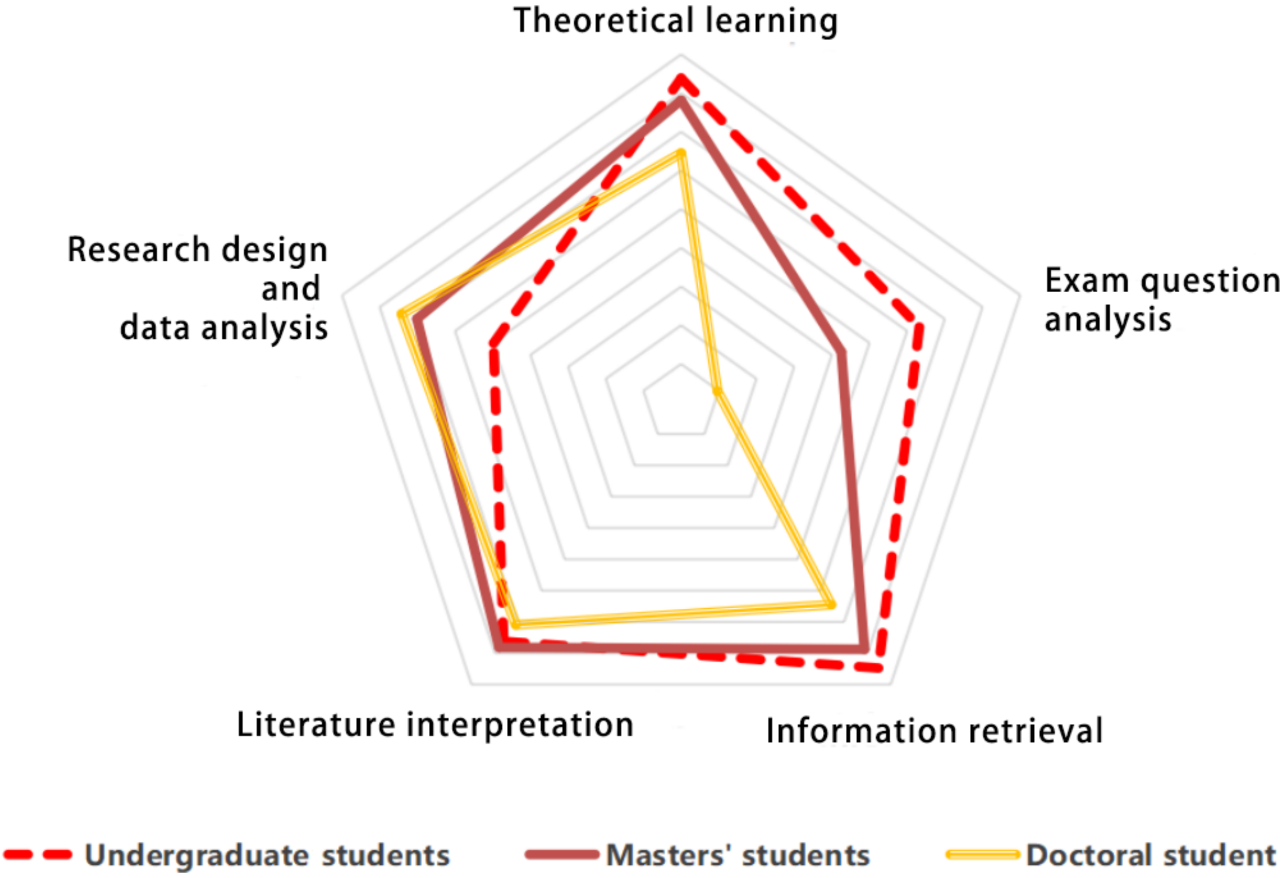
Gender-based differences were observed in the adoption of specific AI platforms, with selected variations reaching statistical significance (χ^2_1=7.84; *P*=.005). Nonetheless, the overall variety of the platforms utilized remained consistent across genders, as evidenced by a comparable number of the tools used (χ^2_7=3.86; *P*=.80).

AI Usage Patterns by Academic Program Type and Educational Stage

The application of AI in medical education spans multiple areas, including theoretical learning, question analysis, literature translation, and scientific research. Based on the literature review and preliminary survey feedback, our questionnaire categorized the application areas into 5 domains: "theoretical learning," "exam question analysis," "information retrieval," "literature interpretation," and "research design and data analysis."

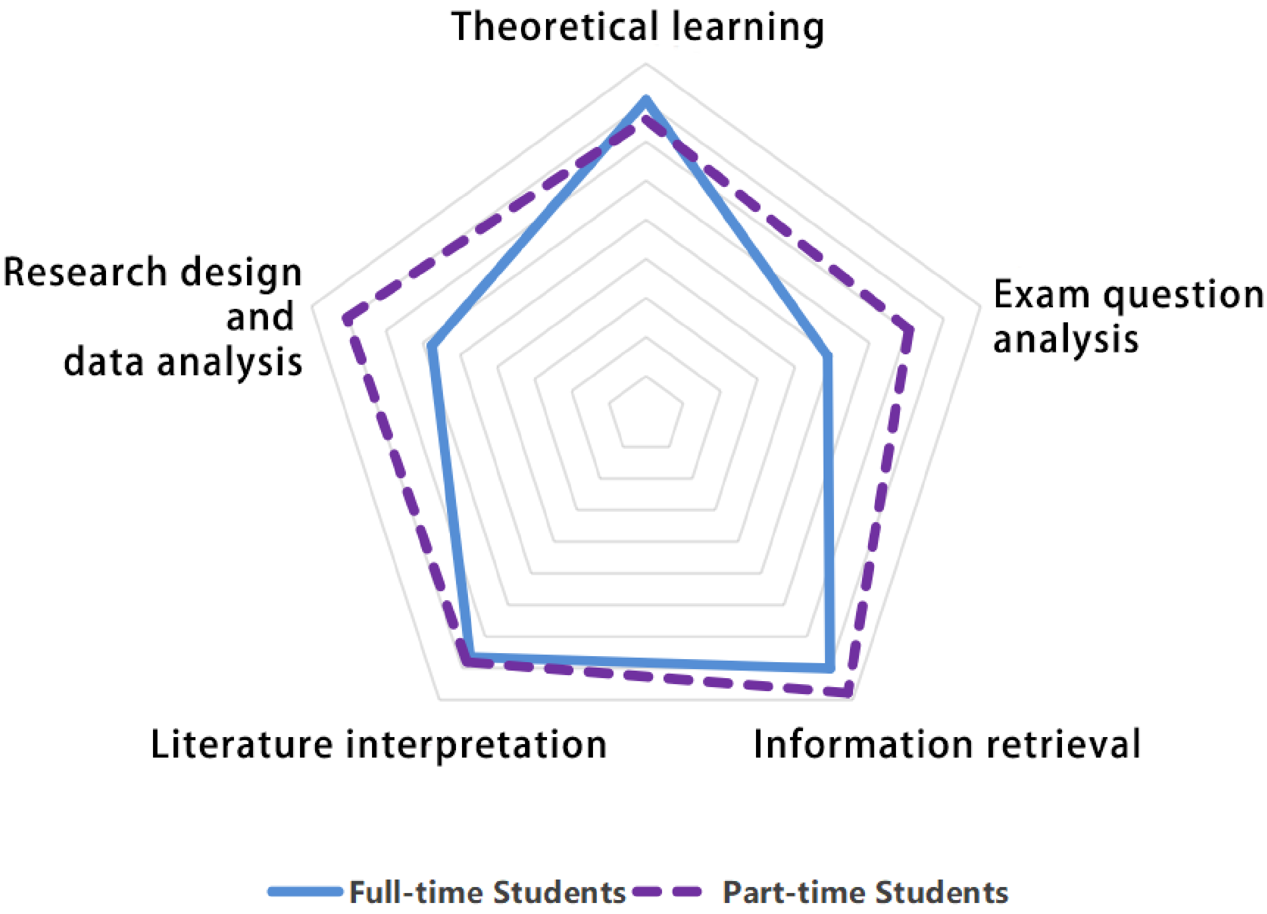
Regarding educational stage, undergraduate students primarily used AI for exam preparation, while graduate students focused more on research tasks such as study design and data analysis (Figure 1).

Figure 1. Usage of artificial intelligence (AI) tools in learning among medical students from different educational stages.



When grouped by academic program type, part-time students demonstrated stronger needs for AI support in practical research and exam question analysis. This preference pattern appears related to their need to balance studies with work commitments, requiring efficient learning solutions that yield immediate academic and professional applicability (Figure 2).

Figure 2. Usage of artificial intelligence (AI) tools in learning among medical students from different academic program types (full-time vs part-time).



AI Usage Patterns by Disciplines

Medical students from different academic disciplines exhibited distinct priorities in their application of AI-powered learning (Table 2). Initial analysis using a multiple response test revealed a significant overall difference in AI usage patterns across disciplines (χ^2_{30}=53.62; *P*=.005), prompting subsequent pairwise comparisons between disciplines with a Bonferroni-adjusted alpha of 0.01.

Table . Comparative analysis of artificial intelligence (AI) function usage across different medical disciplines.

Domain	Clinical medicine (n=245), n (%)	Nursing (n=65), n (%)	Rehabilitation therapy (n=31), n (%)	Basic medicine (n=31), n (%)	Public health and epidemiology and pharmacy[a] (n=11), n (%)	Traditional Chinese medicine (n=22), n (%)	Medical engineering (n=23), n (%)	Chi-square (*df*)	Cramer *V*	*P* value[b]
Theoretical learning	201 (82.04)	50 (76.92)	24 (77.42)	26 (83.87)	8 (72.73)	16 (72.73)	18 (78.26)	2.58 (6)	0.08	.86
Exam question analysis	139 (56.73)	37 (56.92)	12 (38.71)	14 (45.16)	4 (36.36)	7 (31.82)	5 (21.74)	18.49 (6)	0.21	.005
Information retrieval	202 (82.45)	56 (86.15)	19 (61.29)	28 (90.32)	8 (72.73)	15 (68.18)	18 (78.26)	13.89 (6)	0.18	.03
Literature interpretation	184 (75.10)	55 (84.62)	22 (70.97)	26 (83.87)	10 (90.91)	12 (54.55)	19 (82.61)	11.80 (6)	0.17	.07
Research design and data analysis	144 (58.78)	35 (53.85)	17 (54.84)	23 (74.19)	9 (81.82)	15 (68.18)	13 (56.52)	7.02 (6)	0.13	.32

[a] Disciplines including public health and epidemiology and pharmacy were grouped together for analysis due to their relatively small sample sizes.

[b] Chi-squared test and Bonferroni correction were applied for multiple comparisons (α=.01).

Across all disciplines, the usage demand for theoretical learning was similarly high, with no statistically significant differences identified.

Regarding exam question analysis, students in clinical medicine demonstrated significantly higher usage demand than those in medical engineering (χ^2_1=10.36; *P*=.001).

In terms of information retrieval, students from rehabilitation therapy demonstrated significantly lower usage demand compared to students in clinical medicine (χ^2_1=7.72; *P*=.005), nursing (χ^2_1=7.59; *P*=.006), and basic medicine (χ^2_1=7.12; *P*=.008).

For literature interpretation, traditional Chinese medicine students demonstrated a significantly lower level of usage demand relative to students in nursing (χ^2_1=8.40; *P*=.004).

Across disciplines, usage demand for research design and data analysis was moderate, peaking non-significantly among public health and epidemiology and pharmacy students.

Correlates of Satisfaction With the Institutional AI Platform

With the rapid development of AI, many universities have launched localized AI platforms. Our survey investigated the availability of institution-specific AI−powered medical education platforms among medical students. Furthermore, we assessed student satisfaction with these platforms.

Approximately one-fifth (86/428) of the respondents reported that their institutions had developed such platforms. Satisfaction scores among these users exhibited significant variation. The average satisfaction score among the 86 users was 72.23 (SD 21.84), distributed as 40 (46.51%) satisfied, 28 (32.56%) neutral, and 18 (20.93%) dissatisfied.

Nonparametric tests revealed a significant difference in satisfaction levels by gender (*U*=686.50; *z*=−2.06; *P*=.04). No significant associations were found with academic program type, educational stage, or discipline. A subsequent multivariable regression that included these variables and usage frequency identified no significant predictors. To assess the model's reliability, collinearity diagnostics were performed, and they revealed no substantial multicollinearity (Table 3).

Table . Medical students’ satisfaction with the artificial intelligence (AI)−powered medical education platforms at their institutions.

Predictor	Unstandardized coefficient		Standardized coefficient	*P* value	VIF[a]	95% CI for unstandardized coefficients (*B*)
	B	SE	β			
Constant	71.97	13.93	—[b]	<.001	—	44.22 to 99.72
Academic program type (reference: part-time)						
Full-time	−3.18	9.72	−0.04	.74	1.20	−22.55 to 16.19
Gender (reference: female)						
Male	8.52	5.12	0.20	.10	1.11	−1.68 to 18.72
Frequency	−0.39	1.34	−0.03	.77	1.08	−3.06 to 2.28
Discipline (reference: clinical medicine)						
Nursing	5.59	6.85	0.10	.42	1.14	−8.05 to 19.24
Rehabilitation therapy	−4.14	13.76	−0.04	.76	1.08	−31.55 to 23.26
Basic medicine	−0.36	7.38	−0.01	.96	1.11	−15.07 to 14.35
Public health and epidemiology and pharmacy	10.54	12.64	0.10	.41	1.20	−14.64 to 35.71
Medical engineering	−6.18	24.99	−0.03	.81	1.22	−55.96 to 43.61
Educational stage (reference: undergraduate student)						
Master’s student	−2.43	5.97	−0.05	.69	1.33	−14.32 to 9.46
Doctoral student	8.59	10.83	0.10	.43	1.29	−12.99 to 30.18

[a] VIF: variance inflation factor.

[b] Not applicable.

The lack of significant predictors for satisfaction should be interpreted with caution. This result may reflect the inherently subjective and multifaceted nature of satisfaction, which can be influenced by factors beyond the scope of this study. Future research employing longitudinal or mixed methods designs is needed to unravel the complex drivers of user satisfaction.

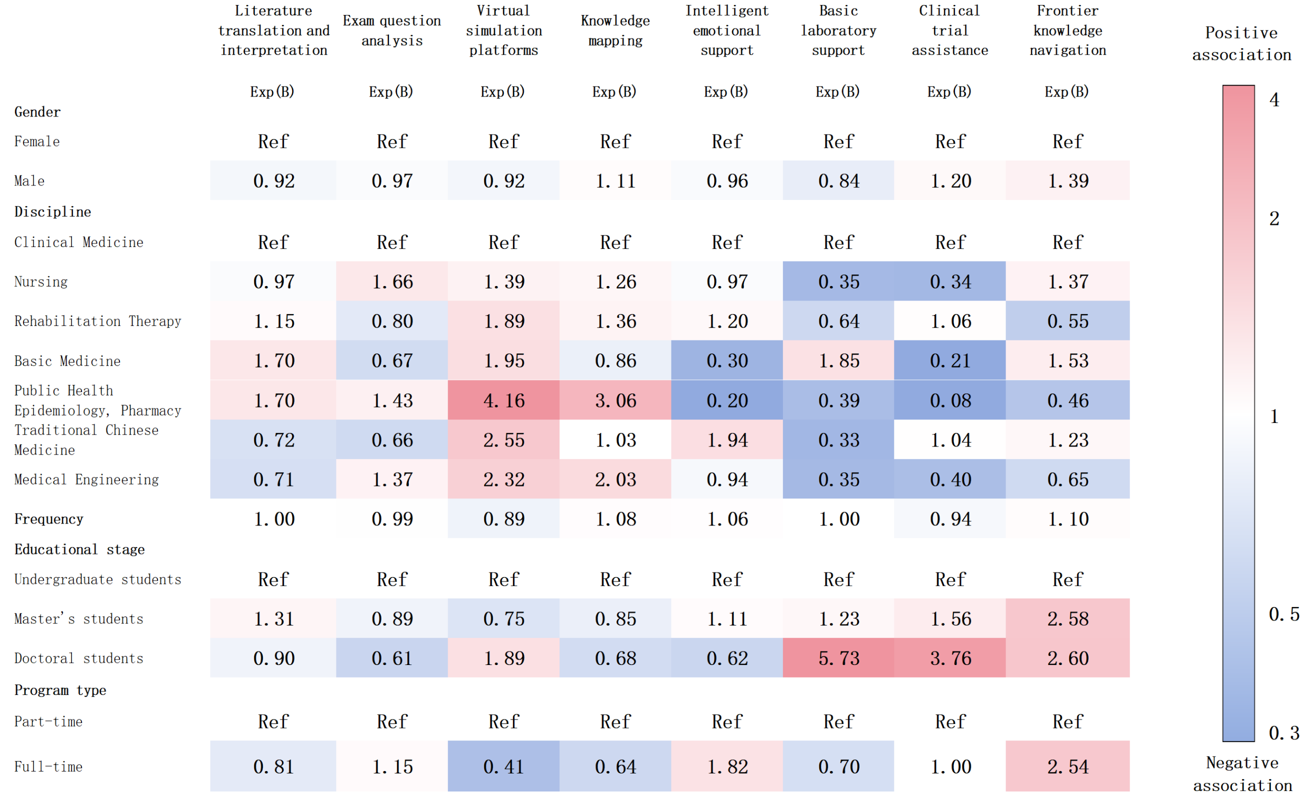
Correlates of Expected Functions for the Institutional AI Platform

The majority (342/428) of the students reported that their institutions had not yet launched an AI platform specifically centered on medical education. To investigate the needs of this group, this survey further explored their practical needs and expectations for upcoming platforms.

Drawing on the current practical applications of AI technology and feedback from medical students, along with a review of relevant literature, our questionnaire categorized the expected functions of AI-powered medical education platforms into 8 aspects: literature translation and interpretation, exam question analysis, clinical trial assistance, basic laboratory support, knowledge mapping, virtual simulation platforms, frontier knowledge navigation, and intelligent emotional support.

Nonparametric testing revealed significant subgroup differences in the demand for specific platform functions. Guided by these initial findings, we advanced the analysis using binary logistic regression within a more rigorous multivariable framework. For comparing the relative effects of multiple predictors, the results are presented as odds ratios (Exp(B)) and visualized in the accompanying heatmap (Figure 3).

Figure 3. Expected functions for artificial intelligence (AI) medical education platforms across student subgroups.



	Literature translation and interpretation	Exam question analysis	Virtual simulation platforms	Knowledge mapping	Intelligent emotional support	Basic laboratory support	Clinical trial assistance	Frontier knowledge navigation
	Exp(B)	Exp(B)	Exp(B)	Exp(B)	Exp(B)	Exp(B)	Exp(B)	Exp(B)
Gender								
Female	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Male	0.92	0.97	0.92	1.11	0.96	0.84	1.20	1.39
Discipline								
Clinical Medicine	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Nursing	0.97	1.66	1.39	1.26	0.97	0.35	0.34	1.37
Rehabilitation Therapy	1.15	0.80	1.89	1.36	1.20	0.64	1.06	0.55
Basic Medicine	1.70	0.67	1.95	0.86	0.30	1.85	0.21	1.53
Public Health Epidemiology, Pharmacy	1.70	1.43	4.16	3.06	0.20	0.39	0.08	0.46
Traditional Chinese Medicine	0.72	0.66	2.55	1.03	1.94	0.33	1.04	1.23
Medical Engineering	0.71	1.37	2.32	2.03	0.94	0.35	0.40	0.65
Frequency	1.00	0.99	0.89	1.08	1.06	1.00	0.94	1.10
Educational stage								
Undergraduate students	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Master's students	1.31	0.89	0.75	0.85	1.11	1.23	1.56	2.58
Doctoral students	0.90	0.61	1.89	0.68	0.62	5.73	3.76	2.60
Program type								
Part-time	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Full-time	0.81	1.15	0.41	0.64	1.82	0.70	1.00	2.54

Regarding gender, the regression analysis revealed a consistent pattern: no statistically significant associations were found between gender and any of the anticipated platform functions.

Analysis of academic program type (full-time vs part-time) revealed no significant differences in functional expectations. However, full-time students demonstrated a numerically stronger preference for frontier knowledge navigation (*B*=0.93; Exp(B)=2.54; *P*=.14), while they were less inclined toward Virtual Simulation Platforms (*B*=−0.90; Exp(B)=0.41; *P*=.08).

Regarding the educational stage, doctoral students demonstrated a significantly stronger preference for AI-assisted clinical trial support (*B*=1.32; Exp(B)=3.76; *P*=.01) and basic laboratory assistance (*B*=1.75; Exp(B)=5.73; *P*=.001) compared to undergraduates, while master's students showed significantly higher demand for frontier knowledge navigation (*B*=0.95; Exp(B)=2.58; *P*=.02).

Across different academic disciplines, medical students showed varying expectations for the functions of upcoming AI-powered medical education platforms. Students in clinical medicine demonstrated a significantly stronger demand for clinical trial assistance (*B*=−1.09; Exp(B)=0.34; *P*=.008) and Basic Laboratory Support (*B*=−1.05; Exp(B)=0.35; *P*=.005) compared to those in nursing. Conversely, nursing students were inclined toward personalized guidance on theoretical exam preparation, though this tendency did not reach statistical significance (*B*=0.51; Exp(B)=1.66; *P*=.08). Notably, students in clinical medicine demonstrated a significantly higher demand for clinical trial assistance compared to students in other disciplines. This difference reached statistical significance when compared specifically with students in basic medicine (*B*=−1.58; Exp(B)=0.21; *P*=.01) and those in public health and epidemiology and pharmacy (*B*=−2.54; Exp(B)=0.08; *P*=.02).

Test of the TTF Hypotheses

Based on the results of our prior data analysis and guided by the 5 research hypotheses derived from the TTF theory, we employed appropriate statistical methods to test these hypotheses (Table 4).

Table . Hypothesis testing results of the task-technology fit (TTF) model.

Hypothesis	Path relationship	Statistic[a]	Effect size[b]	*P* value	Supported
H1	TTF (current use) → satisfaction	$F(5, 80)=1.11$	Adjusted $R^2=0.006$	.36	No
H2	Discipline → TTF (current use)	$F(36, 2526)=1.60$	Partial $\eta^2=0.02$	.01	Yes
	Discipline → TTF (expected function)	$F(42, 840)=1.25$	Partial $\eta^2=0.06$	.13	No
H3	Academic program type → TTF (current use)	$F(6, 421)=3.32$	Partial $\eta^2=0.05$	.003	Yes
	Academic program type → TTF (expected function)	$F(7, 140)=1.22$	Partial $\eta^2=0.06$	.29	No
H4	Educational stage → TTF (current use)	$F(12, 842)=6.51$	Partial $\eta^2=0.09$	<.001	Yes
	Educational stage → TTF (expected function)	$F(14, 280)=1.53$	Partial $\eta^2=0.07$	.10	No
H5	Frequency → satisfaction	$B=-0.39$	95% CI for $B=-3.06$ to 2.28	.77	No

[a] Analytical methods were hypothesis-specific: ANOVA for a single outcome variable (H1); multivariate ANOVA for multiple outcome variables (H2-H4); linear regression with unstandardized coefficients for predictive modeling (H5).

[b] Effect size measures were selected and reported in accordance with the conventions for the specific statistical procedures applied.

The analytical approach was tailored to the characteristics of the variables under examination. For H1, ANOVA was used, which indicated no significant association between current AI usage patterns and satisfaction. For H2, H3, and H4, MANOVA was applied: the results showed that discipline, academic program type, and educational stage each had a significant influence on TTF (current use); however, none of these factors demonstrated a significant effect on TTF (expected function). For H5, linear regression analysis found no significant relationship between frequency and satisfaction.

Discussion

While previous studies have described broad trends in AI adoption among medical students, this study identifies the specific drivers of heterogeneity within a defined cohort from Shanghai. Through a granular subgroup analysis, we demonstrate how disciplinary background, educational stage, and program type significantly shape distinct patterns in usage frequency, functional preferences, and perceived value.

Popularity of AI Technology Among Medical Students

The rapid advancement of AI technology is demonstrating the potential to reshape traditional paradigms in medical education. For medical students, the integration of digital and intelligent technologies has significantly enhanced instructional quality and learning outcomes. With ongoing technological advancement, AI is now widely adopted and has become an essential part of students' academic work [19].

The survey revealed that medical students now engage with AI tools on a frequent basis, reflecting a notable shift in usage patterns compared to earlier adoption phases [20,21]. This suggests that, with ongoing technological advancement and the growing accessibility of AI, students are increasingly adopting this new technology and using it more in their learning.

In terms of LLM selection, most students preferred mainstream models, encompassing both domestic platforms such as DeepSeek and Doubao and international ones like ChatGPT. The choice appears to have been influenced by factors such as accessibility and performance. While mainstream LLMs meet most students' academic requirements, many still select specific models based on personal preferences and practical needs. This reflects a growing trend toward individualized AI tool usage.

Regarding mainstream LLM selection, international findings show that ChatGPT is widely popular globally [22,23]; however, its adoption among the surveyed medical students in Shanghai remains lower compared to DeepSeek. This observed usage pattern aligns with DeepSeek's established localization strategy and regulatory compliance within the Chinese environment. China's regulatory policies require GenAI services to complete local filing and security assessments [24]. ChatGPT, lacking such compliance, is inaccessible through conventional channels. Furthermore, China's exclusion from OpenAI's supported countries creates additional access barriers [25]. In contrast, DeepSeek operates in full compliance with these requirements, ensuring seamless accessibility for Chinese users and thus gaining a competitive edge in the local AI market.

The analysis of AI tool preferences revealed distinct gender−based patterns. Despite using a similar number of AI platforms, with mainstream tools being central for both groups, male students demonstrated greater enthusiasm for emerging options, whereas female students adopted a more cautious approach. These observed differences align with previous findings on gendered perceptions of AI technology [26-28]. Notably, this variation further underscores the importance of developing flexible, multiplatform strategies to accommodate diverse user preferences in medical education.

The integration of AI technology has become a defining feature of contemporary medical education. Our findings reveal a clear consensus among students across all academic backgrounds on adopting a multiplatform approach. Regardless of educational stage or discipline, medical students are actively leveraging diverse AI tools, flexibly selecting platforms according to specific learning scenarios and practical needs [29].

Group Differences in AI Tool Usage for Academic Purposes

Medical knowledge covers a broad spectrum and involves extensive interdisciplinary integration. Although various medical specialties are inherently connected, they differ significantly in teaching method, clinical practice, and scientific research. These distinctions are also reflected in the application of AI.

From an educational stage perspective, undergraduate students tended to emphasize AI-assisted analysis of theoretical exam questions more than master’s and doctoral students, who placed greater focus on AI’s role in supporting practical research. This difference reflects the distinct teaching priorities at each educational stage: undergraduate students face a heavy burden of theoretical courses and professional qualification exams, so their learning focus tends to be on theoretical exam question analysis and knowledge mapping. In contrast, graduate students, facing less exam pressure, focus more on research-related tasks and thus place greater emphasis on AI as a tool for practical research support [30-32].

From the perspective of academic program type (full-time vs part-time), part-time students had broader and more comprehensive needs for AI-assisted learning compared to full-time students. Their demand for AI in scientific research and test analysis was more pronounced. This stems directly from the dual pressures faced by part-time medical students, who balance both academic studies and daily work. In dealing with research and exams, they tend to be more reliant on AI for assistance and guidance to alleviate the burden of learning [33].

From a disciplinary perspective, different majors have a significant impact on the direction of AI-assisted learning [34]. The data showed that students majoring in clinical medicine and nursing were more focused on AI-assisted exam question analysis [35], while traditional Chinese medicine students had a notably lower demand for literature translation and interpretation. Students in basic medicine, public health and epidemiology, and pharmacy were more concerned with AI’s role in supporting research work. These differences directly reflect the distinct curricular focus, professional requirements, and learning objectives characterizing each discipline [7,20,36].

Student-Centered Design of Platform Functions

At present, the application of AI in medical education is still in an exploratory stage. Major computer and internet companies, both domestically and internationally, have launched AI models with diverse functionalities. In parallel, universities and research institutions have adopted various strategies to expand the potential of this emerging field. For example, the AI application platform at Tongji University integrates multiple large models such as DeepSeek, OpenAI, and Tongyi Qianwen, offering a wide range of AI tools. Faculty and students can select these tools based on their professional and academic requirements [37]. Similarly, Shanghai Jiao Tong University has launched the “Jiao Xiao Zhi” agent management platform, integrating LLMs including DeepSeek for localized AI deployment. The platform enables faculty and students to create customized AI assistants, streamlining teaching and research tasks through its secure campus−based infrastructure [38].

This survey focused on medical students’ satisfaction with and expectations for AI-powered medical education platforms. While overall satisfaction was relatively high, a significant minority (approximately one-quarter) expressed dissatisfaction, with notable variations across academic stages and disciplines.

To address these divergent needs, we propose a tiered design framework for future platforms:

For undergraduate students: Platforms should prioritize structured learning support. This includes AI-powered tools for adaptive test preparation aligned with standardized exams, interactive virtual patient cases for foundational clinical exposure, and personalized review systems that target individual knowledge gaps.

For postgraduate students (master’s or PhD): The focus should shift to research and specialized skill development. Key features should encompass advanced literature interpretation aids, data analysis modules for processing experimental or clinical data, and AI assistants for research design and grant writing, catering to their deep engagement in academic research.

Discipline-specific customization: Further refinement should distinguish between clinical and basic medicine disciplines. Clinically oriented tracks would benefit from advanced diagnostic simulators and patient management tools, while basic medicine tracks require robust support for experimental design, genomic data analysis, and scientific visualization.

By moving beyond a one-size-fits-all model to adopt such a stratified and discipline-aware approach, AI platforms can achieve deeper integration into medical education, ultimately enhancing both student satisfaction and educational outcomes [9,39].

Limitations

Due to constraints in personnel and resources, this study employed a convenience sampling method among medical students in Shanghai. This nonprobability sampling approach may have introduced selection bias, such as an overrepresentation of students already enthusiastic about AI, and limits the generalizability of the findings beyond similar urban, well-resourced contexts.

Moreover, this study relied on self-reported measures for key metrics—including AI usage frequency and satisfaction—for which formal psychometric validation was not conducted. This approach increased susceptibility to recall and social desirability biases across all these measures. Furthermore, as subjective constructs, they were vulnerable to varying personal interpretations and benchmarks. This was particularly relevant for abstract constructs like satisfaction. Consequently, these subjective data may not fully capture the nuanced realities of actual user behavior.

Future research could therefore prioritize multicenter, large-sample designs that utilize standardized instruments. Such approaches would provide more robust and generalizable evidence to effectively guide the development of AI-powered medical education platforms.

Conclusions

This study conducted a cross-sectional survey across major medical institutions in Shanghai, China, collecting and analyzing the current use and practical needs of AI among medical students from different educational stages, academic program types, and disciplines. Our findings clearly indicate that AI is widely applied in medical education and has become a common tool for student learning. Students from different disciplines, educational stages, genders, and academic program types show significant differences in their functional demands for AI-assisted learning.

Furthermore, our study investigated the current status of AI-powered medical education platforms and explored students' expectations for such platforms. It is evident that the rapid integration of AI in medical education holds great promise, and our findings provide evidence-based support to guide the future development of AI-powered medical education platforms.

Acknowledgments

We thank all the medical students who participated in this survey for taking time to complete the questionnaires and providing valuable data for this study. We also appreciate those who assisted in questionnaire distribution and collection, which helped improve the response rate and data quality of the research.

We would like to extend special thanks to the Clinical Research Center of Renji Hospital for their guidance in study design, methodology, and ethics, as well as their ongoing technical and logistical support throughout the study.

GenAI was used for language polishing and reference formatting during manuscript preparation. The authors reviewed and edited all content and take the full responsibility for the work.

Funding

This work was funded by National Key Research and Development Program of China (2024YFF0618400), Pudong New Area Health Commission (PWR12024-08), and Yuanshen Rehabilitation Institute Afflicted to Shanghai Jiao Tong University School of Medicine (yskf3-23-1107-2; yskf2-23-1107-2).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy restrictions in the informed consent agreements. However, anonymized data and related documentation are available from the corresponding author upon reasonable request for academic use, subject to a data use agreement and required citation. Requests should include a brief research proposal and will be reviewed within 4 - 6 weeks.

Authors' Contributions

Conceptualization: XS (lead), ZJ (equal)
Formal analysis: XS (lead)
Investigation: XS (lead), LX (supporting)
Methodology: ZJ (lead), LX (supporting), KCS (supporting)
Project administration: ZC (lead)
Supervision: ZC (lead)
Writing – original draft: XS (lead), KCS (supporting)
Writing – review & editing: ZC (lead), XS (supporting), ZJ (supporting), LX (supporting), KCS (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed items of our questionnaire.
[DOCX File, 17 KB - humanfactors_v13i1e81652_app1.docx]

Checklist 1
CHERRIES checklist.

[PDF File, 44 KB - humanfactors_v13i1e81652_app2.pdf]

References

1. Moro C, Štromberga Z, Raikos A, Stirling A. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. Anat Sci Educ 2017 Nov;10(6):549-559. [doi: 10.1002/ase.1696] [Medline: 28419750]
2. Boillat T, Prithishkumar IJ, Suresh D, Naidoo N. Integrating virtual reality to enhance remote teaching of anatomy during unprecedented times. Anat Cell Biol 2025 Mar 31;58(1):112-121. [doi: 10.5115/acb.24.197] [Medline: 39623255]
3. Lee H. The rise of ChatGPT: exploring its potential in medical education. Anat Sci Educ 2024 Dec;17(9):1779. [doi: 10.1002/ase.2496] [Medline: 39087851]
4. Wu Z, Liu Z, Zhang Z. Research on the current situation and future governance of artificial intelligence in the field of medical education. Chin Med Ethics 2024;37(9):1093-1100. [doi: 10.12026/j.issn.1001-8565.2024.09.12]
5. Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2024 Feb 27;10:e57594. [doi: 10.2196/57594] [Medline: 38412478]
6. Wang J, Luo L, Zhao C, Zeng X, Tang R, He S. Practice research of VR-based immersion programs in the clinical emergency. Smart 2022;8(10):176-178. [doi: 10.19335/j.cnki.2096-1219.2022.10.053]
7. Cherrez-Ojeda I, Gallardo-Bastidas JC, Robles-Velasco K, et al. Understanding health care students' perceptions, beliefs, and attitudes toward AI-powered language models: cross-sectional study. JMIR Med Educ 2024 Aug 13;10:e51757. [doi: 10.2196/51757] [Medline: 39137029]
8. Zou L, Lin H, Sun L. Policy text research on AI-enabled medical education. Chin Med Education Technol 2025;39(1):23-28. [doi: 10.13566/j.cnki.cmet.cn61-1317/g4.202501005]
9. Opinions of the state council on deepening the implementation of the "artificial intelligence+" initiative. State Council of the People's Republic of China. 2025. URL: https://cset.georgetown.edu/publication/china-ai-plus-opinions-2025/ [accessed 2025-12-17]
10. Shanghai medical AI development work plan (2025-2027). : General Office of the Shanghai Municipal People's Government; 2024 URL: https://www.shanghai.gov.cn/nw12344/20241224/0f60c01551784720899dbb911e9d5f08.html [accessed 2025-12-12]
11. Zhang J, Li C, Lyu W. The application prospects and challenges of artificial intelligence in medical education, medical research, and clinical practice. Chin J Gen Pract 2024;22(7):1085-1089. [doi: 10.16766/j.cnki.issn.1674-4152.003572]
12. Xiu J, Li Y, Wang Z, et al. Exploration on smart medical education based on 5G technology: taking the construction and application of "5G+Smart Education" in Guizhou Medical University as an example. J Guizhou Med Univ 2024;49(12):1864-1872. [doi: 10.19367/j.cnki.2096-8388.2024.12.021]
13. Tortella F, Palese A, Turolla A, et al. Knowledge and use, perceptions of benefits and limitations of artificial intelligence chatbots among Italian physiotherapy students: a cross-sectional national study. BMC Med Educ 2025 Apr 18;25(1):572. [doi: 10.1186/s12909-025-07176-w] [Medline: 40251635]
14. Tao W, Yang J, Qu X. Utilization of, perceptions on, and intention to use AI chatbots among medical students in China: national cross-sectional study. JMIR Med Educ 2024 Oct 28;10:e57132. [doi: 10.2196/57132] [Medline: 39466038]
15. Lwanga SK, Lemeshow S. Sample size determination in health studies: a practical manual. : World Health Organization; 1991 URL: https://iris.who.int/items/9c2e5da4-3785-4fec-9dbc-841e4ae0d98c [accessed 2025-12-12]
16. Ziapour A, Darabi F, Janjani P, Amani MA, Yıldırım M, Motevaseli S. Factors affecting medical artificial intelligence (AI) readiness among medical students: taking stock and looking forward. BMC Med Educ 2025 Feb 18;25(1):264. [doi: 10.1186/s12909-025-06852-1] [Medline: 39966878]
17. Sami A, Tanveer F, Sajwani K, et al. Medical students' attitudes toward AI in education: perception, effectiveness, and its credibility. BMC Med Educ 2025 Jan 17;25(1):82. [doi: 10.1186/s12909-025-06704-y] [Medline: 39833834]
18. Weidener L, Fischer M. Artificial intelligence in medicine: cross-sectional study among medical students on application, education, and ethical aspects. JMIR Med Educ 2024 Jan 5;10:e51247. [doi: 10.2196/51247] [Medline: 38180787]
19. Ganjavi C, Eppler M, O'Brien D, et al. ChatGPT and large language models (LLMs) awareness and use. A prospective cross-sectional survey of U.S. medical students. PLOS Digit Health 2024 Sep;3(9):e0000596. [doi: 10.1371/journal.pdig.0000596] [Medline: 39236008]
20. Alharbi MK, Syed W, Innab A, Basil A Al-Rawi M, Alsadoun A, Bashatah A. Healthcare students attitudes opinions perceptions and perceived obstacles regarding ChatGPT in Saudi Arabia: a survey-based cross-sectional study. Sci Rep 2024 Oct 1;14(1):22800. [doi: 10.1038/s41598-024-73359-y] [Medline: 39354005]
21. Pan G, Ni J. A cross sectional investigation of ChatGPT-like large language models application among medical students in China. BMC Med Educ 2024 Aug 23;24(1):908. [doi: 10.1186/s12909-024-05871-8] [Medline: 39180023]
22. Ajalo E, Mukunya D, Nantale R, et al. Widespread use of ChatGPT and other artificial intelligence tools among medical students in Uganda: a cross-sectional study. PLoS ONE 2025;20(1):e0313776. [doi: 10.1371/journal.pone.0313776] [Medline: 39787055]
23. von Garrel J, Mayer J. Artificial intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. Humanit Soc Sci Commun 2023;10(1):799. [doi: 10.1057/s41599-023-02304-7]

24. Interim measures for the management of generative artificial intelligence services. : China Cyberspace Administration; 2023 URL: https://openresearch-repository.anu.edu.au/server/api/core/bitstreams/7577e6a0-db26-4dab-9273-8d13a6883a40/content [accessed 2025-12-12]
25. ChatGPT supported countries. OpenAI. URL: https://help.openai.com/en/articles/7947663-chatgpt-supported-countries-and-territories [accessed 2025-12-12]
26. Duan S, Liu C, Rong T, Zhao Y, Liu B. Integrating AI in medical education: a comprehensive study of medical students' attitudes, concerns, and behavioral intentions. BMC Med Educ 2025 Apr 23;25(1):599. [doi: 10.1186/s12909-025-07177-9] [Medline: 40269824]
27. Zhai J, Zhang H. Critical thinking disposition of medical students in Anhui Province, China: a cross-sectional investigation. BMC Med Educ 2023 Sep 8;23(1):652. [doi: 10.1186/s12909-023-04646-x] [Medline: 37684599]
28. Abdelhafiz AS, Farghly MI, Sultan EA, Abouelmagd ME, Ashmawy Y, Elsebaie EH. Medical students and ChatGPT: analyzing attitudes, practices, and academic perceptions. BMC Med Educ 2025 Feb 5;25(1):187. [doi: 10.1186/s12909-025-06731-9] [Medline: 39910593]
29. George Pallivathukal R, Kyaw Soe HH, Donald PM, Samson RS, Hj Ismail AR. ChatGPT for academic purposes: survey among undergraduate healthcare students in Malaysia. Cureus 2024 Jan;16(1):e53032. [doi: 10.7759/cureus.53032] [Medline: 38410331]
30. Xu S, Su T, Wang H, Xu J, Liu P, Tang Y. On the differences and reasons of online courses demand for military medical students at different educational levels: taking an example of a military medical university. Chin J Med Educ Res 2021;20(11):1334-1338. [doi: 10.3760/cma.j.cn116021-20200630-00681]
31. Zhao W, Zhu B. Learn to learn, learn to study, learn to create: different focuses for different levels of students. Acad Degree Graduate Education 2011(2):52-56. [doi: 10.3969/j.issn.1001-960X.2011.02.013]
32. You S, Zhou J, Li J, Kong M, Tan W. Practice and reflection on cultivating top-notch innovative talents in life sciences and medicine. Bull Chin Acad Sci 2025;40(2):260-272. [doi: 10.16418/j.issn.1000-3045.20241106001]
33. Alsobhi M, Khan F, Chevidikunnan MF, Basuodan R, Shawli L, Neamatallah Z. Physical therapists' knowledge and attitudes regarding artificial intelligence applications in health care and rehabilitation: cross-sectional study. J Med Internet Res 2022 Oct 20;24(10):e39565. [doi: 10.2196/39565] [Medline: 36264614]
34. Li Y, Xu J, Jia C, Zhai X. Investigation of college students' generative artificial intelligence (GAI) usage status and its implication: taking Zhejiang University as an example. Open Educ Res (1):89-98. [doi: 10.13966/j.cnki.kfjyyj.2024.01.010]
35. Wu N, Chen S, Yu B, Chen X. The current situation and reflection on the application of generative artificial intelligence assisted learning for vocational nursing students. Chin J Nurs Educ (4):403-409. [doi: 10.3761/j.issn.1672-9234.2025.04.004] [Medline: 24063287]
36. Liu S, Wright AP, Patterson BL, et al. Using AI-generated suggestions from ChatGPT to optimize clinical decision support. J Am Med Inform Assoc 2023 Jun 20;30(7):1237-1245. [doi: 10.1093/jamia/ocad072] [Medline: 37087108]
37. DeepSeek Tongji localized model user guide. Sohu. URL: https://www.sohu.com/a/862666794_712343 [accessed 2025-10-12]
38. Jiao Xiao Zhi Agent Management Platform. Xi'an Jiaotong University. URL: https://agent.xjtu.edu.cn/index.htm [accessed 2025-10-12]
39. Elhassan SE, Sajid MR, Syed AM, Fathima SA, Khan BS, Tamim H. Assessing familiarity, usage patterns, and attitudes of medical students toward ChatGPT and other chat-based AI apps in medical education: cross-sectional questionnaire study. JMIR Med Educ 2025 Jan 30;11:e63065. [doi: 10.2196/63065] [Medline: 39883912]

Abbreviations

AI: artificial intelligence
LLM: large language model
TTF: task-technology fit model

Edited by A Kushniruk; submitted 31.Jul.2025; peer-reviewed by AR Hoque, B Mohan, F Anupama, P Usuemerai, R Mamidipally; revised version received 02.Dec.2025; accepted 04.Dec.2025; published 08.Jan.2026.

Please cite as:
Shi X, Jiang Z, Xiong L, Siu KC, Chen Z
Utilization of AI Among Medical Students and Development of AI Education Platforms in Medical Institutions: Cross-Sectional Study
JMIR Hum Factors 2026;13:e81652
URL: https://humanfactors.jmir.org/2026/1/e81652
doi:10.2196/81652

Parental Influence on Children’s Media Use in South Korea: National Population-Based Study

Ji Young Kim[1], PhD; Ah Jung Yang[2], PhD; Hye Eun Lee[3], PhD

[1]Department of Psychology, School of Behavioral Sciences and Education, Pennsylvania State University, Middletown, PA, United States

[2]Communication & Media Research Center, College of Social Science, Ewha Womans University, Seoul, Republic of Korea

[3]Department of Communication & Media, College of Social Science, Ewha Womans University, #403 Ewha-POSCO, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul, Republic of Korea

Corresponding Author:
Hye Eun Lee, PhD
Department of Communication & Media, College of Social Science, Ewha Womans University, #403 Ewha-POSCO, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: To better understand the effects of media use on children, it is essential to examine the various factors influencing the media use of digital native children. In the situational context, parental media usage, parents’ attitudes toward media, and parenting styles have all been identified as significant factors influencing children’s media use. This study focuses on the key factors and examines these relationships in greater depth, drawing on existing research to understand their impact on the media usage patterns of digital native children.

Objective: This study examines parental influences related to young children’s media use in Korea over a 3-year period (2022 - 2024) using independent, nationally representative cohorts.

Methods: Using multigroup structural equation modeling, we analyzed data from 3 independent parent–reported cohorts (for 2022, n=1058; for 2023, n=1020; for 2024, n=1020) to investigate how parental media habits, attitudes, and distinct parenting styles predict children’s daytime and nighttime media consumption.

Results: The online survey results revealed that parental media time, particularly for mothers, consistently correlated with higher levels of children’s daytime media use (β=.002-.003). Positive parental attitudes toward media increased children’s daytime media use (β=.028-.102), whereas negative attitudes had a limited effect (β=−.069-.140). Among the 7 parenting styles, positive parenting consistently reduced children’s daytime media use in 2022 and 2023 (β=−.228 for 2022, β=−.215 for 2023), but harsh punishment emerged as the strongest factor in daytime media use in 2024 (β=−.078 for 2022, β=−.090 for 2023, and β=−.072 for 2024). Notably, parenting styles showed no significant effect on children’s nighttime media use throughout the study, suggesting that parental influence may be more effective during daytime hours.

Conclusions: This analysis extends existing research by differentiating media use patterns across time periods and highlights the evolving influence of parenting styles. These findings have implications for the development of targeted parental guidelines for managing young children’s media exposure, especially as digital media continues to become a pervasive part of daily life.

(JMIR Hum Factors 2026;13:e75292) doi:10.2196/75292

KEYWORDS

children media use; parent attitude; parent media use; parental mediation; parenting style

Introduction

Parental Influences on Children’s Media Use in South Korea

An increasing number of children access media at younger ages, with many actively consuming it. In the United States, 70% of children under the age of 1 year and 91% of children aged 2-3 years access media several times per week [1]. Similarly, in South Korea, approximately 60% of children aged 3-9 years are exposed to television before their first birthday, and 30% begin using smartphones before the age of 1 year [2]. This growing trend is in contrast with the recommendation by the American Academy of Pediatrics [3], which advises against media use for children under the age of 2 years [4]. Despite the recommendation, children’s media consumption remains high. The typical 8- to 10-year-old children spend almost 8 hours daily across different media, and this number increases for older children and teenagers to more than 11 hours per day [1].

While media may not be a leading cause of major health concerns, previous research suggests that media exposure contributes to various risks and health problems [3]. For example, younger children’s media use, as early as preschool years, has been associated with developmental difficulties such as reduced physical fitness and psychosocial health [5,6]. Specifically, engagement in risky online social behaviors [7], poor academic performance [8], obesity [9,10], reduced quality of life [11], and impaired executive function [12] have been associated with increased children's media use. Additionally, previous studies have shown that excessive television viewing can lead to cognitive, language, and social-emotional delays [13-16]. Research has further shown that infants and toddlers gain more developmentally beneficial skills through play with physical objects rather than screen-based devices [17,18].

Numerous longitudinal [19] and cross-sectional studies [5,13,20] showed that children with greater screen time had a higher level of behavioral difficulties. A longitudinal study investigating children between the ages of 2 and 6 years showed an association between increased media use and poorer well-being outcomes, even though the results varied for boys and girls [13]. Additionally, each extra hour of television viewing per week between the ages of 2 and 4 years was associated with a 5% increase in body mass index at the age of 10 years, a 9%‐10% rise in the consumption of soda and snacks, and a 13% reduction in outdoor weekend activities [21]. A study from South Korea with survey data from mothers and teachers of 5- and 6-year-old children showed that increased use of smart devices was associated with increased levels of aggression [22]. Altogether, the findings suggest that children's excessive consumption of media may have adverse developmental and behavioral effects.

To better understand the effects of media use on children, it is essential to examine the various factors influencing the media use of digital native children (ie, children who have access to and utilize a wide range of media and platforms from birth) [23]. One key factor is the physical environment, or in other words, the advancements in digital media technology that enable children to use media freely without restrictions of time and space through traditional means such as television as well as portable devices like smartphones and tablet PCs. Another key factor is the situational context, such as the influence of parents, given that parents shape their children's behaviors and attitudes as primary role models [24-26]. Parental media usage [27-29], attitudes toward media [30,31], and parenting styles [26,32] have all been identified as significant factors influencing children's media use. In addition, the impact of these factors varies according to the developmental stage or age of the children [25,26,32]. This study seeks to focus on the second key factor and examine these relationships in greater depth, drawing on existing research to understand their impact on the media usage patterns of digital native children.

Although extensive research has examined children's overall media or screen time, much less is known about how parental factors influence media use across different times of day. This is a critical gap because daytime and nighttime media use may have distinct developmental implications (eg, school readiness, sleep quality, and behavioral regulation). Moreover, little evidence exists on how these influences shift across post-COVID years when children's reliance on screen-based media has been both normalized and expanded to include educational contexts [33,34]. This study addresses this problem by examining parental media use, attitudes, and parenting styles as predictors of children's daytime and nighttime media use in a large, national Korean sample drawn from three independent cohorts across 3 years (2022‐2024). In this study, "media use" is defined operationally as children's screen-based engagement, including television, smartphones, tablets, and computers, as reported by parents in a 24-hour matrix of daily activities. Conceptually, we focus on electronic and online media rather than print or offline formats, as these account for the majority of children's daily exposure in contemporary contexts.

Parents' Media Use and Social Learning Theory

Children's screen time can be explained by Bandura's social learning theory [35], which posits that learning and behavior are shaped by observing others in one's environment. Young children learn by observing their parents, siblings, and others around them, picking up on daily habits, social interactions, and responses to various situations, including media use in the home. When parents model media behaviors, their children are likely to replicate those patterns. For instance, a national survey found that children often imitate their fathers or older siblings by playing with game controllers, even when they lack the skills to use them effectively [36].

Parents' prolonged or habitual consumption of media at home may serve as a role model for children's media use. Research indicates that parents' media use and habitual patterns of consumption are associated with increased media use time and habitual use among children [37]. Notably, children's media use was more closely related to their parents' habitual consumption of media than to parental restrictions on media use. Further, studies have shown that children's screen time tends to be higher when their mothers report substantial media use [38,39]. These findings suggest that parents' media use may significantly influence children's ability to regulate their own media consumption. Accordingly, we propose:

- H1: A parent's screen time will be positively related to a child's media use.

Parent Attitudes and Media

Parent attitudes are another important factor in children's screen time. Parents play a critical role in establishing rules within the household, with their influence on media consumption being particularly pronounced in younger children. Parent attitudes toward media, whether positive or negative, shape the rules governing media use, thereby directly influencing the amount of time children spend engaging with media [26]. Given their substantial influence on their children's behaviors [40], parents often modify their behaviors as a strategy to affect changes in their children's behaviors [41].

In the United States, a survey focused on electronic media use among children aged 6 months to 6 years indicated that parents had varied attitudes toward media use [36]. Some parents encouraged media use, finding it beneficial for children and useful for household management, such as keeping children occupied while doing chores. While similar proportions of parents felt that television had both positive (38%) and negative (31%) effects on learning, the majority supported computers as beneficial (70%) but video games as detrimental (49%) [29]. It is possible that parents aim to help their children maximize the benefits of the rich resources offered by video media while simultaneously shielding them from harmful content [42].

In general, the children of parents who view media positively tend to have higher screen time. Among preschoolers, parent attitudes and beliefs about media were strong indicators of children's screen time [30]. For younger (0 - 2 y) and older (5 - 6 y) children, positive parent attitudes have been shown to significantly contribute to exceeding the recommended screen time limits set by the American Academy of Pediatrics [3,43], particularly in relation to television viewing [44]. Although parent attitudes significantly influenced the television and computer use of children, their influence on mobile devices was less pronounced [31].

Additionally, parental rules based on positive attitudes tend to support media use in the home, while negative attitudes encourage restrictions. The more critically parents evaluate media, the more likely they are to engage positively in parental mediation (active or restrictive) to mitigate the negative effects of their children's media exposure [41]. Parental mediation has been shown to effectively influence children's responses to media exposure [45], and experts have even advocated for increasing the level of parental mediation provided by parents [46]. Therefore, we hypothesize:

- H2: Parents with a positive attitude toward media will have children with higher levels of media use.
- H3: Parents with a negative attitude toward media will have children with lower levels of media use.

Parenting Style and Media Use

Finally, parents considerably contribute to children's media use, indicating that parents may mediate children's media use and, in turn, the overall developmental outcome. For example, while children with increased media use showed a decreased level of prosocial behavior, the level of parent-child interactions mediated the association between the children's media use and level of prosocial behavior [19]. Specifically, parenting style—or specific sets of parental mediation strategies and behaviors—may explain children's varying levels of media use. Previous studies have shown that parental control over children's access to smart media influences the amount of time younger children spend using such media [47]. Research in Korea has further indicated that among parental mediation strategies, supervisory control tends to be more pronounced before the age of 6 years but diminishes after the age of 7 years. In particular, for children aged 5-6 years, multiple factors—including media use patterns, self-regulatory abilities, and parental supervision—exert significant influence. This suggests that age-specific approaches to parents' mediation are necessary [48]. Research also indicates that mediation practices vary not only in form (eg, restrictive, active, or co-use) but also in effectiveness, depending on the child's age. For younger children, restrictive mediation (eg, setting time limits) has been shown to reduce overall screen exposure, whereas older children and adolescents benefit more from active mediation approaches that involve discussion, explanation, and joint engagement [25,32,49-51].

Baumrind [52] and Maccoby and Martin [53] identified four different parenting styles: authoritative, authoritarian, indulgent, and neglectful. Each parenting style varied in their levels of demandingness and responsiveness. Generally, studies showed that children with authoritative parents (ie, highly demanding and responsive) had the most positive developmental outcome, while children with authoritarian (ie, highly demanding but not responsive) and permissive (ie, not demanding but highly responsive) parents had more negative outcomes.

The Ghent Parental Behavior Scale, developed by Leeuwen and Vermulst [54], categorizes parenting strategies into 9 distinct domains: positive parenting, autonomy, rules, monitoring, discipline, harsh punishment, material reward, inconsistent discipline, and ignoring. A material reward is a tangible incentive, such as a toy, snack, small gift, or even cash, provided to a child in exchange for a specific behavior or compliance. Positive parenting styles, such as permissive and neglectful parenting, are characterized by strategies including autonomy, positive parenting, and rules. In contrast, negative parenting styles, such as authoritative or authoritarian approaches, are associated with parenting strategies including monitoring, harsh punishment, material rewards, and ignoring [54].

A recent study by Lee et al [26] showed that a child's daytime media use between the ages of 4 and 6 years significantly increased for parents who gave more autonomy to their children (ie, permissive and neglectful parenting style), and a child's nighttime media use decreased with the parenting style of discipline (ie, authoritative and authoritarian parenting style). Lee et al also found that parents' use of material rewards was the strongest predictor of the child's nighttime media use, as the child's nighttime media use significantly increased when parents practiced material reward. Similarly, other studies showed that children between the ages of 10 and 11 years exhibited higher levels of screen exposure when they had parents with lower levels of control over their children [27,32]. The parental behavior in this example aligns with a permissive and neglectful parenting style, where parents provide more autonomy to their children and incorporate positive parenting. On the contrary, studies showed that parents who exercise higher levels of control and support would engage in behaviors or strategies employing active and restrictive mediation [32,55]. Specific strategies would include monitoring and rules, discipline, harsh punishment, and material reward. Collectively, the literature suggests that children with permissive and neglectful parents would be more likely to have higher levels of media use, while children with authoritarian and authoritative parents would be more likely to have lower levels of media use [26].

Despite the clear implications of the parent's role in shaping children's media use, previous studies have focused on descriptive and correlational analyses. As Lee et al [26] suggested, further investigation on the relationship between parental determinants, such as parenting style and children's media use, is needed. The primary goal of the study is to determine the relationship between parenting style and children's media use across various age groups. The researcher distributed surveys to parents because young children cannot reliably self-report their daily media use, and parent reports are generally considered appropriate for estimating children's media use, such as video and app games, computer use, and television viewing [56]. Based on previous findings that parenting style may potentially have important effects on children's media use, we hypothesize the following:

- H4: Specific parenting styles will be associated with children's level of media use.

The secondary goal of the study is to extend the findings of Lee et al [26] by replicating the previous findings. Therefore, we incorporated the 2 hypotheses related to parenting style used in Lee et al [26]:

- H5: Children who have parents with permissive and neglectful parenting styles (ie, exercise positive parenting and give autonomy to children) will have higher levels of media use.
- H6: Children who have parents with authoritative and authoritarian parenting styles (ie, monitoring, rules, discipline, harsh punishment, and material reward) will have lower levels of media use.

Previous findings support a clear relationship between parenting style and children's media use. To extend the existing literature, it is imperative to identify the parental determinant that explains children's media use, and the researcher of this study aims to answer the following research question:

- RQ1: Among parent's media use, media attitudes, and parenting styles, what is the most influential factor on children's media use?

Methods

Participants

A total of 1058 parents of children aged 5-7 years, 1020 parents of children aged 6-8 years, and 1020 parents of children aged 7-9 years completed questionnaires between April 26 and May 9, 2022; October 26 and November 9, 2023; and July 31 and August 21, 2024, respectively. All participants completed the questionnaire through an online survey conducted by a Korean survey company, Macromill Embrain, which recruited participants from its national panel pool. Each year's sample primarily consisted of an independent cross-sectional group, although approximately 5% - 10% of the participants may have overlapped across years due to the nature of the panel system. However, the study did not track individuals over time, and analyses were conducted separately for each dataset. The age and sex distribution of the children was balanced across each dataset. Detailed demographic characteristics for all three groups are presented in Table 1.

Table . Demographic characteristics of participants and their child (age, gender, education level, income, employment, and family members).

Year and demographic characteristics	Participants, n (%)
2022	
Age (child) (y)	
5	359 (33.9)
6	352 (33.3)
7	347 (32.8)
Gender (child)	
Male	535 (50.6)
Female	523 (49.4)
Age (parent) (y)	
20	15 (1.4)
30	578 (54.6)
40	453 (42.8)
50	12 (1.1)
Gender (parent)	
Male	129 (12.2)
Female	929 (87.8)
Family members	
2	11 (1.0)
3	301 (28.4)
4	576 (54.4)
5	109 (10.3)
>6	61 (5.8)
Education level (father)	
High school or below	144 (13.6)
College or university	781 (73.8)
Master’s degree or above	133 (12.6)
Education level (mother)	
High school or below	108 (10.2)
College or university	834 (58.8)
Master’s degree or above	116 (11.0)
Monthly income (family) (USD)	
<4000	341 (32.2)
4000 - 5000	235 (22.2)
5000 - 6000	161 (15.2)
>6000	321 (30.3)
Employment (father)	
Full-time job	993 (93.9)
Part-time job	50 (4.7)
No job	15 (1.4)
Employment (mother)	
Full-time job	519 (49.1)
Part-time job	171 (16.2)

Year and demographic characteristics	Participants, n (%)
No job	368 (34.8)
Total	1058 (100)
2023	
Age (child) (y)	
6	340 (33.3)
7	340 (33.3)
8	340 (33.3)
Gender (child)	
Male	510 (50.0)
Female	510 (50.0)
Age (parent) (y)	
20	5 (0.5)
30	468 (45.9)
40	541 (53.0)
50	6 (0.6)
Gender (parent)	
Male	144 (14.1)
Female	876 (85.9)
Family members	
2	317 (31.1)
3	547 (53.6)
4	105 (10.3)
5	31 (3.0)
>6	20 (2.0)
Education level (father)	
High school or below	128 (12.6)
College or university	759 (74.4)
Master’s degree or above	133 (13.0)
Education level (mother)	
High school or below	90 (8.8)
College or university	803 (78.7)
Master’s degree or above	127 (12.5)
Monthly income (family) (USD)	
<4000	205 (20.1)
4000 - 5000	222 (21.8)
5000 - 6000	173 (16.9)
>6000	420 (41.2)
Employment (father)	
Full-time job	968 (94.9)
Part-time job	41 (4.0)
No job	11 (1.1)
Employment (mother)	
Full-time job	576 (56.5)

Year and demographic characteristics	Participants, n (%)
Part-time job	160 (15.7)
No job	284 (27.8)
Total	1020 (100)
2024	
Age (child) (y)	
7	340 (33.3)
8	340 (33.3)
9	340 (33.3)
Gender (child)	
Male	510 (50.0)
Female	510 (50.0)
Age (parent) (y)	
20	7 (0.7)
30	369 (36.2)
40	628 (61.6)
50	16 (1.6)
Gender (parent)	
Male	133 (13.0)
Female	887 (87.0)
Family members	
2	307 (31.1)
3	549 (53.8)
4	119 (11.7)
5	33 (3.2)
>6	12 (1.2)
Education level (father)	
High school or below	123 (12.1)
College or university	741 (72.6)
Master’s degree or above	156 (15.3)
Education level (mother)	
High school or below	82 (8.0)
College or university	810 (79.5)
Master’s degree or above	128 (12.5)
Monthly income (family) (USD)	
<4000	172 (16.9)
4000 - 5000	197 (19.3)
5000 - 6000	179 (17.5)
>6000	472 (46.3)
Employment (father)	
Full-time job	976 (95.7)
Part-time job	40 (3.9)
No job	4 (0.4)
Employment (mother)	

Year and demographic characteristics	Participants, n (%)
Full-time job	602 (59.0)
Part-time job	153 (15.0)
No job	265 (26.0)
Total	1020 (100)

Ethical Considerations

The study received ethical approval from the Institutional Review Board of Ewha Womans University (Institutional Review Board approval number: EWHA-202103-0028-01). All data were anonymized prior to analysis, and no personally identifiable information was stored or retained. Informed consent was obtained from all participants involved in the study. As the participants were members of a registered survey company panel, an honorarium was provided based on the survey response time at a rate of 100 KRW (approximately US $0.67) per minute. Participants who discontinued the survey prior to completion were also compensated with a prorated stipend corresponding to their participation time.

Instrument and Measures

Overview

The questionnaire was initially developed in English and then translated into Korean, with translation equivalence verified by bilingual researchers. In addition to the main variables, media use (ie, time spent on media via television, personal device, smartphone) and demographic information of both children and their parents were collected. Tables 2, 3, and 4 present the reliabilities and descriptive statistics of the variables, along with their correlations. Composite variables were calculated once unidimensionality and acceptable reliability were confirmed. All variables were measured with 5-point Likert scales (1="strongly disagree" to 5="strongly agree") unless otherwise noted.

Table . Reliabilities, correlations, means, and standard deviations of the main variables (2022)[a].

2022	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Child’s age at first use	__[b]														
Child’s locus of control	–0.05	(0.75)													
Mother’s media time	0.01	0.17[c]	—												
Father’s media time	0.04	0.13[c]	0.62[c]	—											
Positive attitude toward media use	0.10[c]	–0.02	0.10[c]	0.05	(0.90)										
Negative attitude toward media use	0.01	0.44[c]	0.03	0.07[d]	–0.26[c]	(0.89)									
Positive parenting	–0.05	–0.19[c]	–0.06	–0.03	–0.09[c]	–0.12[c]	(0.89)								
Monitoring	0.01	–0.06[d]	–0.07[d]	–0.02	0.10[c]	0.02	0.48[c]	(0.72)							
Rules	–0.07[d]	–0.11[c]	0.01	0.02	0.01	–0.01	0.52[c]	0.44[c]	(0.84)						
Discipline	0.00	0.13[c]	0.02	0.04	0.05	0.21[c]	–0.05	0.10[c]	0.21[c]	(0.76)					
Harsh punishment	0.05	0.26[c]	0.11[c]	0.07[d]	0.10[c]	0.22[c]	–0.36[c]	–0.16[c]	–0.20[c]	0.34[c]	(0.92)				
Material reward	0.05	0.25[c]	0.11[c]	0.10[c]	0.21[c]	0.12[c]	–0.04	0.07[d]	0.00	0.21[c]	0.25[c]	(0.76)			
Autonomy	0.05	–0.10[c]	0.01	0.06	0.13[c]	–0.03	0.45[c]	0.33[c]	0.37[c]	0.03	–0.13	0.07[d]	(0.80)		
Child’s daytime media use	–0.08[d]	0.34[c]	0.26[c]	0.17[c]	0.04	0.10[c]	–0.12[c]	–0.07[d]	–0.06	0.05	0.05	0.08[c]	–0.04		—
Child’s nighttime media use	–0.01	0.12[d]	0.12[c]	0.11[c]	–0.05	0.02	–0.05	–0.06	–0.04	0.05	0.08[c]	0.01	–0.03	0.09**[c]	

2022	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mean (SD; range)	3.33 (1.36; 1 - 8)	2.88 (0.82; 1 - 5)	120.67 (83.81; 0 - 600)	127.10 (86.03; 0 - 600)	2.92 (0.70; 1 - 5)	3.09 (0.68; 1 - 5)	4.01 (0.47; 1.91 - 5)	3.79 (0.60; 1 - 5)	4.12 (0.46; 1.86 - 5)	3.31 (0.66; 1 - 5)	1.71 (0.86; 1 - 4.75)	2.83 (0.84; 1 - 5)	3.83 (0.53; 2 - 5)	2.37 (1.49; 0 - 13)	0.19 (0.48; 0 - 4.29)

[a] Reliabilities, calculated using Cronbach α, are reported in parentheses on the diagonal.

[b] Not applicable.

[c] *P*<.001.

[d] *P*<.05.

Table . Reliabilities, correlations, means, and standard deviations of the main variables (2023)[a].

2023	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Child’s age at first use	—[b]														
Child’s locus of control	–0.04	(0.76)													
Mother’s media time	0.01	0.16[c]	—												
Father’s media time	–0.03	0.15[c]	0.59[c]	—											
Positive attitude toward media use	0.11[c]	–0.05	–0.02	–0.08[c]	(0.92)										
Negative attitude toward media use	–0.01	0.42	0.02	0.05	–0.22[c]	(0.90)									
Positive parenting	–0.03	–0.22[c]	–0.13[c]	–0.07[d]	0.15[c]	–0.11[c]	(0.9)								
Monitoring	–0.03	–0.05	–0.05	0.01	0.08[c]	0.05	0.50[c]	(0.69)							
Rules	–0.07[d]	–0.11[c]	–0.02	0.02	0.01	–0.01	0.54[c]	0.49[c]	(0.85)						
Discipline	–0.04	0.13[d]	0.02	0.04	0.07[d]	0.13[c]	0.04	0.23[c]	0.26[c]	(0.74)					
Harsh punishment	0.06	0.25[c]	0.12[c]	0.09[c]	0.16[c]	0.13[c]	–0.35[c]	–0.16[c]	–0.25[c]	0.25[c]	(0.92)				
Material reward	0.06	0.26[c]	0.07[d]	0.03	0.13[c]	0.16[c]	–0.08[d]	–0.00	–0.06[d]	0.17[c]	0.25[c]	(0.77)			
Autonomy	0.01	–0.13[c]	–0.04	–0.03	0.07[d]	–0.05	0.47[c]	0.31[c]	0.38[c]	0.11[c]	–0.15[c]	0.02	(0.81)		
Child’s daytime media use	–0.10[c]	0.32[c]	0.25[c]	0.22[c]	–0.05	0.10[c]	–0.10[c]	–0.05	–0.01	0.04	0.02	0.05	–0.03		—
Child’s nighttime media use	–0.01	0.15[c]	0.11[c]	0.10[c]	–0.03	0.04	–0.08	–0.04	–0.04	0.06	0.05	0.07[d]	–0.01	0.16[c]	

2023	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mean (SD; range)	3.46 (1.50; 1 - 9)	2.83 (0.84; 1 - 5)	117.53 (81.51; 0 - 610)	125.65 (83.92; 0 - 600)	2.85 (0.72; 1 - 4.89)	3.16 (0.67; 1 - 5)	4.04 (0.48; 2.09 - 5)	3.90 (0.57; 1.75 - 5)	4.32 (0.50; 2.33 - 5)	3.50 (0.72; 1 - 5)	1.70 (0.83; 1 - 4.75)	2.79 (0.83; 1 - 5)	3.86 (0.52; 2 - 5)	2.35 (1.48; 0 - 10.71)	0.22 (0.46; 0 - 4)

[a] Reliabilities, calculated using Cronbach α, are reported in parentheses on the diagonal.

[b] Not applicable.

[c] *P*<.001.

[d] *P*<.05.

Table . Reliabilities, correlations, means, and standard deviations of the main variables (2024)[a].

2024	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Child's age at first use	—[b]														
Child's locus of control	–0.06	(0.76)													
Mother's media time	0.02	0.13[c]	—												
Father's media time	–0.06[d]	0.08[c]	0.54[c]	—											
Positive attitude toward media use	0.04	0.02	0.04	–0.03	(0.93)										
Negative attitude toward media use	–0.01	0.23[c]	0.01	0.01	–0.32[c]	(0.90)									
Positive parenting	–0.01	–0.11[c]	–0.11[c]	–0.07[d]	0.08[d]	–0.08[d]	(0.87)								
Monitoring	–0.02	–0.04	–0.02	0.02	0.03	0.05	0.47	(0.72)							
Rules	–0.05	–0.06	–0.02	0.04	–0.08[d]	0.03	0.54	0.45[c]	(0.92)						
Discipline	–0.02	0.10[c]	0.07[d]	0.06[d]	0.07[d]	0.18[c]	0.06	0.15[c]	0.14[c]	(0.72)					
Harsh punishment	0.02	0.11[c]	0.12[c]	0.09[c]	0.14[c]	0.12[c]	–0.29[c]	–0.14[c]	–0.24[c]	0.37[c]	(0.92)				
Material reward	0.04	0.15[c]	0.03	0.02	0.17[c]	0.14[c]	–0.17[c]	0.01	–0.14[c]	0.23[c]	0.37[c]	(0.77)			
Autonomy	0.04	–0.04	–0.06	–0.04	0.06	0.02	0.40[c]	0.22[c]	0.32[c]	0.01	–0.11	0.01	(0.80)		
Child's daytime media use	–0.13[c]	–0.23[c]	0.18[c]	0.14[c]	–0.05	0.05	–0.05	–0.05	–0.01	–0.02	0.02	0.02	0.03	–0.02	—
Child's nighttime media use	–0.08[d]	0.13[c]	0.15[c]	0.13[c]	0.04	0.01	–0.01	–0.01	0.00	0.04	0.02	0.02	0.06[d]	0.02	0.23[c]

2024	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Mean (SD; range)	3.73 (1.69; 1 - 10)	3.01 (0.61; 1 - 5)	124.52 (85.88; 0 - 600)	140.77 (87.96; 0 - 600)	2.54 (0.79; 1 - 4.89)	3.32 (0.67; 1 - 5)	4.03 (0.46; 2.27 - 5)	3.94 (0.57; 1.75 - 5)	4.43 (0.52; 2.33 - 5)	3.21 (0.60; 1 - 5)	1.73 (0.87; 1 - 5)	2.60 (0.85; 1 - 5)	3.87 (0.57; 1 - 5)	3.48 (2.50; 0 - 14)	0.34 (0.74; 0 - 10)

[a] Reliabilities, calculated using Cronbach α, are reported in parentheses on the diagonal.

[b] Not applicable.

[c] $P<.001$.

[d] $P<.05$.

Children's Age at First Media Use

Parents were asked about their child's age at first media exposure with the question, "How old was your child when he/she first started watching media content?" The response options included "less than 12 months," "1 year," "2 years," "3 years," "4 years," "5 years," "6 years," and "7 years."

Parents' Media Time

Each participant reported both their own and their spouse's media watching time. Two questions were used: "How many hours and minutes do you spend watching media on a typical weekday?" and "How many hours and minutes do you spend to watch media on a typical weekend?" Responses were averaged after weighing (ie, (weekday time × 5 days + weekend × 2 days)/7)) and converted to minutes.

Parents' Positive and Negative Attitudes Toward Media Use

Based on scales by Elias and Sulkin [57] and Nikken and Jansz [58], 9 items measured positive attitudes (eg, "I think watching media will positively influence my child's behavioral development."). Negative attitudes were measured with 2 dimensions: intellectual and social. Each dimension included 4 items, such as "I think watching media will hurt my child's creativity." for the intellectual dimension and "I believe watching media will negatively affect my child's play with friends." for the social dimension.

Parenting Styles

The Ghent Parental Behavior Scale [54] was used to measure the parenting styles. This scale has demonstrated valid structure in diverse samples. Seven dimensions were included due to reliability concerns with 2 original dimensions (inconsistent discipline and ignoring). The included dimensions are positive parenting (11 items: eg, "I make time to listen to my child when he/she wants to tell me something."), monitoring (4 items: eg, "I keep track of the friends my child is seeing."), rules (6 items: eg, "I teach my child to obey rules."), discipline (4 items: eg, "When my child does something I don't approve of, I punish him/her."), harsh punishment (4 items: eg, "I spank my child when he/she is disobedient."), material reward (3 items: eg, "I reward my child with money or a small gift for good behavior."), and autonomy (3 items: eg, "I teach my child to solve problems independently."). Items excluded for reliabilities included 1 ("I teach my child respect for the authorities.") from the rules dimension and 2 items ("When my child has been misbehaving, I give him/her a chore for punishment." and "It happens that I don't punish my child after he/she has done something that is not allowed.") from harsh punishment dimension. All items and dimensions can be referenced in Leeuwen and Vermulst's measurement study [54].

Child's Media Time

Participants selected a cell in a 24-hour matrix to indicate when a child watched media during both weekdays and weekends. Daytime use was defined as viewing between 7:00 AM and 9:00 PM, while viewing before 7:00 AM or after 9:00 PM was classified as nighttime use. This classification was based on developmental sleep guidelines: given that elementary schools in Korea typically begin at 9:00 AM, and the American Academy of Sleep Medicine [59] recommends 10 - 13 hours of sleep for children aged 3 - 5 and 9 - 12 hours for children aged 6 - 12, we estimated that 10 - 11 hours of sleep starting from 9:00 PM would allow children to function healthily on a daily basis. In practical terms, a 9:00 PM bedtime enables children to sleep until approximately 7:00-8:00 AM, providing sufficient time to prepare for and arrive at school. In addition, "screen time" was measured during 2022 - 2024, when academic activities had resumed their normal patterns following the COVID-19 pandemic, thereby minimizing the potential impact of external variables. Media time was averaged and reported in hours. Due to the varied and skewed nature of the data, standardized *z* scores were used for analysis.

Results

Multigroup structural equation modeling was conducted to test the hypotheses and research question using Mplus 8.0 [60], which uses the maximum likelihood estimation method. To evaluate the model fit, confirmatory fit index (CFI), goodness-of-fit index (GFI), and normed fit index (NFI) were used.

Acceptable goodness-of-fit indices were obtained for the overall model [61]: χ^2_1=1.2, P=.27, CFI=0.99, GFI=1.00, NFI=0.99 for 2022, χ^2_1=8.9, P<.01, CFI=0.99, GFI=1.00, NFI=0.99 for 2023, and χ^2_1=44.7, P<.001, CFI=0.98, GFI=0.99, NFI=0.98 for 2024. The estimated coefficients are presented in Figures 1-3.

Figure 1. The final model of the relationships among a child's daytime and nighttime media use, parents' media time, parents' attitude toward media use, and parenting style in 2022. Note that the values are the observed standardized path coefficients. Solid lines indicate significant coefficients at *P*<.05, whereas dotted lines represent nonsignificant coefficients. In addition, paths leading to nighttime media use are shown in light blue, while paths leading to daytime media use are shown in black.

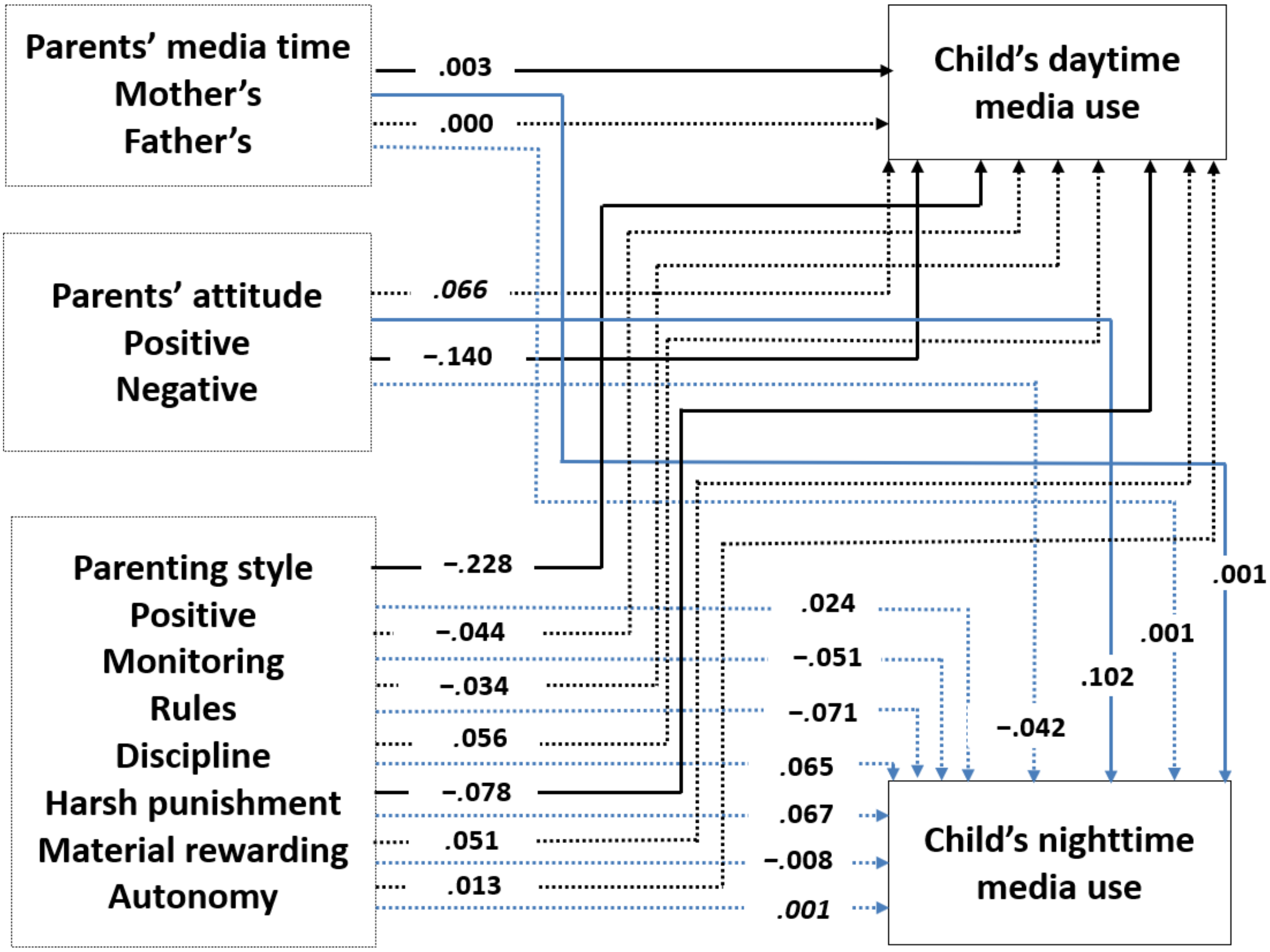


Figure 2. The final model of the relationships among a child's daytime and nighttime media use, parents' media time, parents' attitude towards media use, and parenting style in 2023. Note that the values are the observed standardized path coefficients. Solid lines indicate significant coefficients at *P*<.05, whereas dotted lines represent non-significant coefficients. In addition, paths leading to nighttime media use are shown in light blue, while paths leading to daytime media use are shown in black.

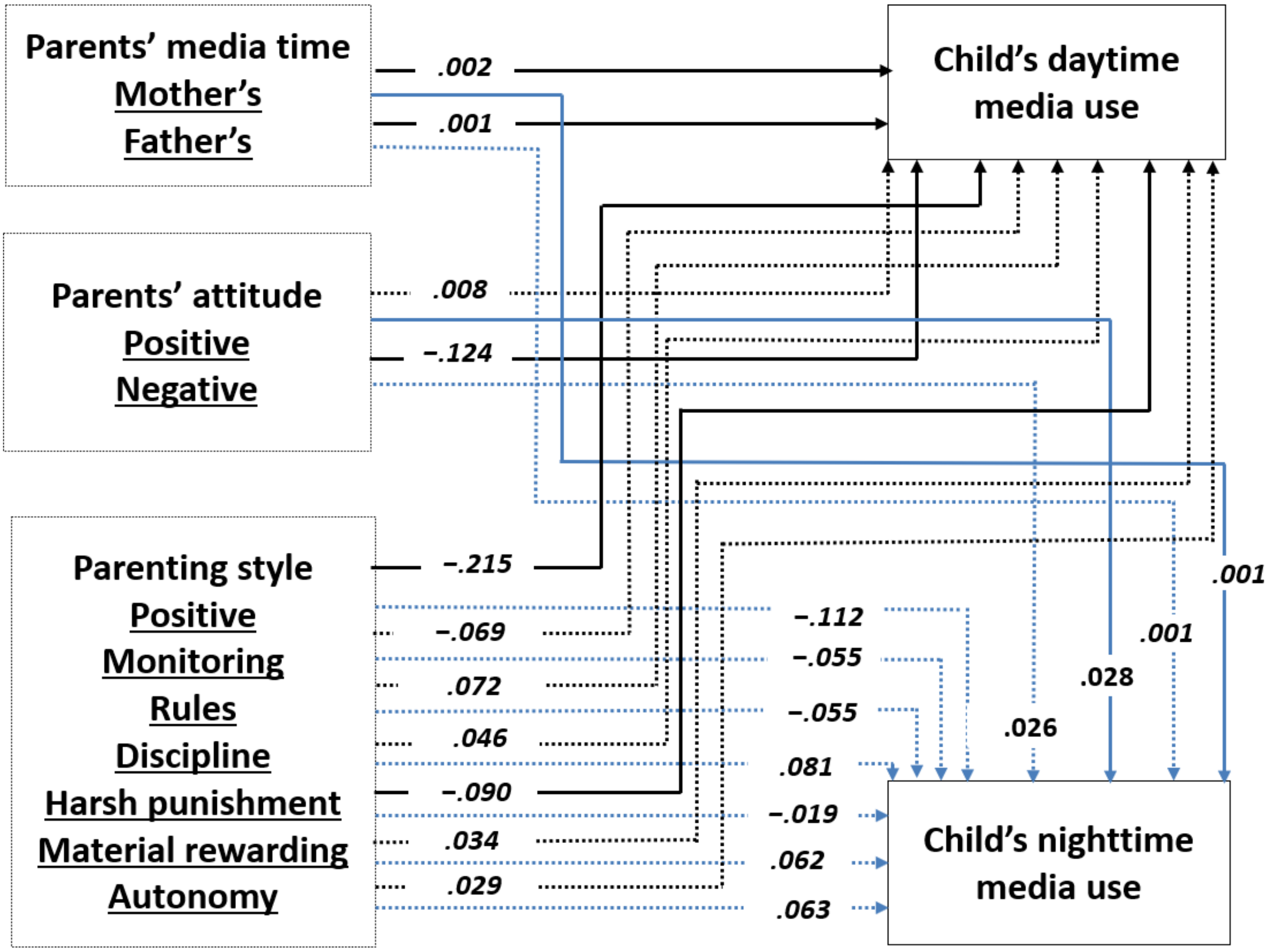
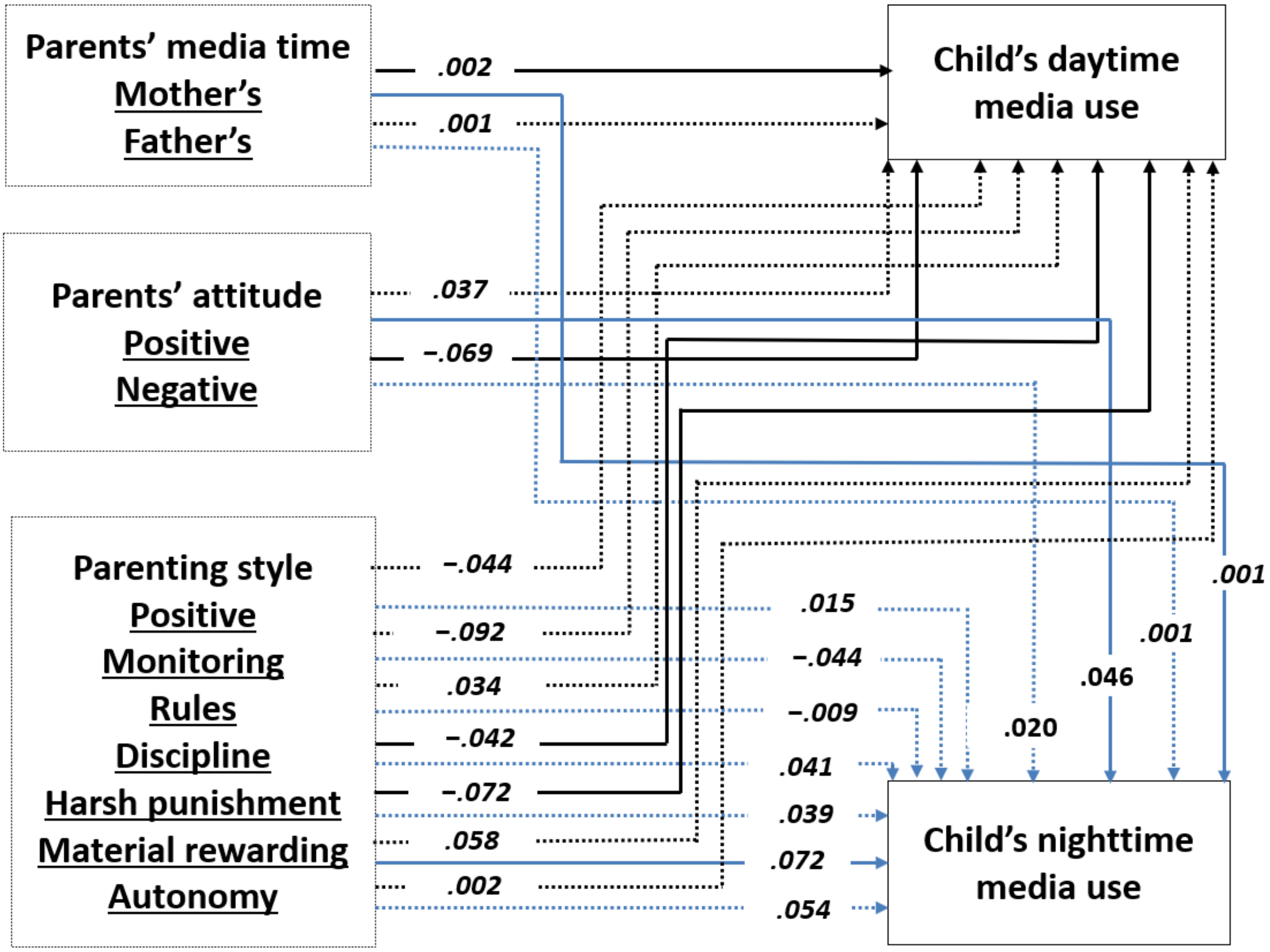


Figure 3. The final model of the relationships among a child’s daytime and nighttime media use, parents’ media time, parents’ attitude toward media use, and parenting style in 2024. Note that the values are the observed standardized path coefficients. Solid lines indicate significant coefficients at *P*<.05, whereas dotted lines represent nonsignificant coefficients. In addition, paths leading to nighttime media use are shown in light blue, while paths leading to daytime media use are shown in black.



H1 predicted that parents with a higher level of screen time would have children who have a higher level of media use. The results partially supported this hypothesis. Regarding a child’s daytime media use, mother’s media time had a positive effect consistently in 2022, 2023, and 2024 (β=.002-.003), but father’s media time showed a significant positive effect on child’s daytime in 2023 only (β=.001). Regarding a child’s nighttime media use, mother’s media time showed a positive effect in 2022, 2023, and 2024 (β=.001).

H2 and H3 hypothesized that parents’ positive and/or negative attitudes toward media use would affect child’s media use. Parents’ positive attitude toward media use increased a child’s nighttime media use (β=.028-.102) and parents’ negative attitude decreased a child’s daytime media use (β=−.069-.140). H2 and H3 were supported.

Finally, H4 predicted that parents with permissive and neglectful parenting styles would be positively associated with children’s media use, while parents with authoritative and authoritarian parenting styles would be negatively related to child’s media use. Positive parenting style decreased a child’s daytime media use (β=−.228 for 2022, β=−.215 for 2023), and harsh punishment (β=−.078 for 2022, β=−.090 for 2023, and β=−.072 for 2024) among the 7 parenting styles significantly decreased a child’s daytime media use. Material reward was positively associated with child’s nighttime media use (β=.072) in 2024.

RQ1 examined which factors were the most influential among parents’ media use, attitude on media, and parenting styles. The results showed that parents’ positive parenting style was the strongest predictor of a child’s daytime media use in 2022 and 2023, but parents’ harsh punishment was the strongest predictor in 2024. Parents’ positive and negative attitudes on media use were the strongest predictor of a child’s nighttime media use in 2022 only.

Discussion

This study investigated the connection between young children’s media use and various parental influences, including media time, attitudes toward media, and parenting styles. The results across 2022, 2023, and 2024 consistently demonstrate the critical role parents play in shaping their children’s media use habits. Similar to previous studies, it was observed that when parents—especially mothers—engage in higher media use, their children’s media consumption during both daytime and nighttime increases, supporting the findings from past research [26,38,39]. This aligns with the idea that parental behavior directly influences children’s media use habits.

Prior research has typically collapsed media consumption into a single metric, overlooking potential temporal differences. Our findings show that parental influences such as parent attitudes and parenting styles are differentially affected by the temporal contexts (ie, daytime vs nighttime). Specifically looking at parent attitudes, across the 3 years, a negative parent attitude toward media consistently correlated with decreased daytime media use, but a positive parent attitude toward media consistently correlated with increased nighttime media use. These consistent relationships across 3 years reflect the findings by Cingel and Krcmar [30] that emphasize the predictive power of parental attitudes on children's screen time.

Parenting style showed distinct influences on children's media use patterns over the 3-year period. In 2022 and 2023, positive parenting emerged as a significant factor in reducing children's daytime media use, highlighting that nurturing and supportive approaches were effective in moderating screen time during the day. By 2024, however, harsh punishment had become the strongest predictor of reduced daytime media use, suggesting a shift where more stringent disciplinary actions had a notable impact on limiting screen time. This change may reflect an evolving role of disciplinary measures in influencing children's media habits over time.

Contrary to expectations, none of the parenting styles showed a significant association with children's nighttime media use across the 3 years except 1. In 2024, material reward was positively associated with a child's nighttime media use. The findings indicated that parenting styles, while impactful on daytime media consumption, may have a limited role in controlling screen time during nighttime hours. Overall, the findings of daytime and nighttime media use suggest that while positive parenting consistently influences daytime media use, the emergence of harsh punishment and material reward as a stronger factor in 2024 points to potential shifts in the effectiveness of various parenting strategies over time.

The findings from a study focused on a younger age group (aged 3-6 y) than this study [26] demonstrated that children's daytime media use increased when parents provided more autonomy to their children. Conversely, nighttime media use decreased among children of parents who emphasized discipline. Children's nighttime media use was significantly increased when parents employed material rewards as a motivational tool. This study extended the findings of Lee et al [26] by investigating an older group of children (ie, aged 5-7 y). The findings from the 2 studies collectively suggest that as children grow older, the effectiveness of these parental interventions, particularly in reducing nighttime media use, diminishes significantly. These findings contribute to a broader understanding of the role of parental determinants in shaping children's media use and underscore the necessity for age-specific parental guidelines.

Previous results consistently highlighted the value of parental mediation in amplifying the positive effects and mitigating the adverse effects of children's media exposure [45,62]. These findings have encouraged scholars to highlight the critical role of parental engagement in managing children's media use [41]. However, a significant gap persists in the implementation of parental controls. For instance, most American adolescents report having no rules or restrictions from their parents regarding either the type of television content or the amount of time spent viewing [1].

Parents play a crucial role in ensuring safe and beneficial use of digital devices, particularly given that children's cognitive and functional abilities are still in developmental stages. Studies have shown that parents can actively promote children's growth and development through well-informed media management practices. Rather than relying on restrictive approaches, Wu et al [63] proposed integrating restrictive, instructive, and joint approach strategies to create a balanced and effective framework. This combined approach encourages collaboration and communication between parents and children, facilitating the establishment of developmentally appropriate media use habits that support overall growth.

Furthermore, these results extend previous research on parental determinants of children's media use and underscore the importance of establishing specific parental guidelines tailored to different times of day. The differentiation between daytime and nighttime media use is particularly relevant, as studies have shown that media use affects multiple developmental domains, including sleep quality, cognitive development, language, and socioemotional and physical health [13-16]. The variations observed over the years suggest that a longitudinal approach can provide a more comprehensive understanding of parental influences on media habits.

From a policy perspective, these findings support the creation of targeted recommendations that help parents foster healthy media use patterns at home. With children increasingly exposed to media, guidelines that encourage informed parental choices could mitigate the potential risks associated with excessive screen time. This evidence-based approach can empower parents to optimize on- and off-screen activities for their children's development.

However, this study is not without limitations. First, it is important to acknowledge that the data were collected during COVID-19 and post-COVID years (2022 - 2024), when screen time became both more prevalent and more socially normalized. In this period, children's daytime media use may not solely reflect leisure activities but may also encompass school-related or educational screen exposure. Second, parental self-reports remain susceptible to social desirability bias, especially given the awareness of recommended screen time limits [64]. Future research should incorporate observational methods to provide a more objective measure of media use. Additionally, distinguishing between types of media would allow for a more nuanced understanding of the effects of specific content on children's behavior and development, as different media types (eg, gaming vs educational content) may have distinct impacts. Similarly, a limitation of this study is the reliance on parent-reported 24-hour matrix data, which, while useful for distinguishing daytime and nighttime use, does not capture the content, context, or quality of children's media exposure. Prior work highlights that screen time measures often overlook these dimensions and that more comprehensive approaches are needed to assess family media environments [65,66]. Future studies should integrate both time-based metrics and contextual measures to provide a fuller understanding of children’s media use. Although the survey company’s online panel is designed to mirror the Korean population in terms of key demographic characteristics (eg, sex, age, and family structure), the sampling method is not entirely random. Therefore, the findings should not be interpreted as being representative of the whole population. Rather, the results should be considered in the context of the study’s sample, and caution should be exercised when generalizing these patterns to all Korean parents and children. Finally, the correlational nature of this study does not imply causation; thus, further research should explore potential genetic and environmental moderators that could influence the relationship between parental behavior and children’s media habits.

In conclusion, the study across 3 years underscores those parental influences, including media time, attitudes, and specific parenting strategies, continue to shape children’s media behaviors. By accounting for temporal patterns, researchers and policymakers can provide parents with tailored guidance to support balanced media use, enhancing children’s overall well-being in the digital age.

Funding

This work was supported by the Ministry of Education of the Republic of the Korea and National Research Foundation of Korea (NRF-2020S1A3A2A02095619).

Conflicts of Interest

None declared.

References

1. Rideout VJ, Foehr UG, Roberts DF. Generation M2: media in the lives of 8- to 18-year-olds. : Kaiser Family Foundation; 2010 URL: https://www.kff.org/wp-content/uploads/2013/04/8010.pdf [accessed 2026-01-09]
2. Children media use survey [Report in Korean]. : Korean Press Foundation; 2020 URL: https://www.kpf.or.kr/front/research/consumerDetail.do?seq=591511 [accessed 2026-01-09]
3. American Academy of Pediatrics, Council on Communications and Media. Children, adolescents, and the media. Pediatrics 2013 Nov;132(5):958-961. [doi: 10.1542/peds.2013-2656] [Medline: 28448255]
4. Lammers SM, Woods RJ, Brotherson SE, Deal JE, Platt CA. Explaining adherence to American Academy of Pediatrics screen time recommendations with caregiver awareness and parental motivation factors: mixed methods study. JMIR Pediatr Parent 2022 Apr 5;5(2):e29102. [doi: 10.2196/29102] [Medline: 35380541]
5. Poulain T, Vogel M, Neef M, et al. Reciprocal associations between electronic media use and behavioral difficulties in preschoolers. Int J Environ Res Public Health 2018 Apr 21;15(4):814. [doi: 10.3390/ijerph15040814] [Medline: 29690498]
6. LeBlanc AG, Spence JC, Carson V, et al. Systematic review of sedentary behaviour and health indicators in the early years (aged 0-4 years). Appl Physiol Nutr Metab 2012 Aug;37(4):753-772. [doi: 10.1139/h2012-063] [Medline: 22765839]
7. Pillay S, Vijaindren A. Keep the smartphones away from your kids. New Straits Times. 2016. URL: http://www.nst.com.my/news/2016/05/147114/keep-smartphonesaway-your-kids [accessed 2026-01-09]
8. Burns MK, Pulles SM, Maki KE, et al. Accuracy of student performance while reading leveled books rated at their instructional level by a reading inventory. J Sch Psychol 2015 Dec;53(6):437-445. [doi: 10.1016/j.jsp.2015.09.003] [Medline: 26563597]
9. Hagen EW, Starke SJ, Peppard PE. The association between sleep duration and leptin, ghrelin, and adiponectin among children and adolescents. Curr Sleep Medicine Rep 2015 Dec;1(4):185-194. [doi: 10.1007/s40675-015-0025-9]
10. Miller DP. Maternal work and child overweight and obesity: the importance of timing. J Fam Econ Iss 2011 Jun;32(2):204-218. [doi: 10.1007/s10834-010-9244-x]
11. Palmer S. Toxic Childhood: How the Modern World Is Damaging Our Children and What We Can Do about It: Orion; 2005. URL: https://openlibrary.org/books/OL40293153M/Toxic_Childhood [accessed 2026-01-09]
12. Huber B, Yeates M, Meyer D, Fleckhammer L, Kaufman J. The effects of screen media content on young children’s executive functioning. J Exp Child Psychol 2018 Jun;170:72-85. [doi: 10.1016/j.jecp.2018.01.006] [Medline: 29448235]
13. Hinkley T, Verbestel V, Ahrens W, et al. Early childhood electronic media use as a predictor of poorer well-being: a prospective cohort study. JAMA Pediatr 2014 May;168(5):485-492. [doi: 10.1001/jamapediatrics.2014.94] [Medline: 24639016]
14. Reid Chassiakos YL, Radesky J, Christakis D, Moreno MA, Cross C, Council on Communications and Media. Children and adolescents and digital media. Pediatrics 2016 Nov;138(5):e20162593. [doi: 10.1542/peds.2016-2593] [Medline: 27940795]
15. Tomopoulos S, Dreyer BP, Berkule S, Fierman AH, Brockmeyer C, Mendelsohn AL. Infant media exposure and toddler development. Arch Pediatr Adolesc Med 2010 Dec;164(12):1105-1111. [doi: 10.1001/archpediatrics.2010.235] [Medline: 21135338]

16. Zimmerman FJ, Christakis DA, Meltzoff AN. Associations between media viewing and language development in children under age 2 years. J Pediatr 2007 Oct;151(4):364-368. [doi: 10.1016/j.jpeds.2007.04.071] [Medline: 17889070]
17. Brown A, Smolenaers E. Parents' interpretations of screen time recommendations for children younger than 2 years. J Fam Issues 2018 Jan;39(2):406-429. [doi: 10.1177/0192513X16646595]
18. Ginsburg KR, American Academy of Pediatrics Committee on Communications, American Academy of Pediatrics Committee on Psychosocial Aspects of Child and Family Health, American Academy of Pediatrics Committee on Psychosocial Aspects of Child and Family Health. The importance of play in promoting healthy child development and maintaining strong parent-child bonds. Pediatrics 2007 Jan;119(1):182-191. [doi: 10.1542/peds.2006-2697] [Medline: 17200287]
19. Zhao J, Zhang Y, Jiang F, et al. Excessive screen time and psychosocial well-being: the mediating role of body mass index, sleep duration, and parent-child interaction. J Pediatr 2018 Nov;202:157-162.e1. [doi: 10.1016/j.jpeds.2018.06.029] [Medline: 30100232]
20. Byun YH, Ha M, Kwon HJ, et al. Mobile phone use, blood lead levels, and attention deficit hyperactivity symptoms in children: a longitudinal study. PLoS One 2013;8(3):e59742. [doi: 10.1371/journal.pone.0059742] [Medline: 23555766]
21. Pagani LS, Fitzpatrick C, Barnett TA, Dubow E. Prospective associations between early childhood television exposure and academic, psychosocial, and physical well-being by middle childhood. Arch Pediatr Adolesc Med 2010 May;164(5):425-431. [doi: 10.1001/archpediatrics.2010.50] [Medline: 20439793]
22. Kim B, Shin H. The relationship between young children's media use time, self-regulation, and problem behaviors. Korean J Hum Ecol 2019 Dec 31;28(6):567-581. [doi: 10.5934/kjhe.2019.28.6.567]
23. Margaryan A, Littlejohn A, Vojt G. Are digital natives a myth or reality? University students' use of digital technologies. Comput Educ 2011 Feb;56(2):429-440. [doi: 10.1016/j.compedu.2010.09.004]
24. An SM, Kang MJ. The effect of mothers' inappropriate smart device usage habits on preschoolers' cognitive ability: sequential mediating effect of preschoolers' smart device immersion tendency and self-regulation. Fam Environ Res 2021 May;59(2):157-168. [doi: 10.6115/fer.2021.012]
25. Collier KM, Coyne SM, Rasmussen EE, et al. Does parental mediation of media influence child outcomes? A meta-analysis on media time, aggression, substance use, and sexual behavior. Dev Psychol 2016 May;52(5):798-812. [doi: 10.1037/dev0000108] [Medline: 26914217]
26. Lee HE, Kim JY, Kim C. The influence of parent media use, parent attitude on media, and parenting style on children's media use. Children (Basel) 2022 Jan 1;9(1):37. [doi: 10.3390/children9010037] [Medline: 35053662]
27. Jago R, Davison KK, Thompson JL, Page AS, Brockman R, Fox KR. Parental sedentary restriction, maternal parenting style, and television viewing among 10- to 11-year-olds. Pediatrics 2011 Sep;128(3):e572-e578. [doi: 10.1542/peds.2010-3664] [Medline: 21859910]
28. Park YM, Kang MJ. Pathway from maternal parenting efficacy, inappropriate motives for allowing smart devices, and smart device dependency to preschoolers' ability to understand minds. Korean J Child Stud 2020 Dec;41(6):9-24. [doi: 10.5723/kjcs.2020.41.6.9]
29. Xu H, Wen LM, Rissel C. Associations of maternal influences with outdoor play and screen time of two-year-olds: findings from the Healthy Beginnings Trial. J Paediatr Child Health 2014 Sep;50(9):680-686. [doi: 10.1111/jpc.12604] [Medline: 24893682]
30. Cingel DP, Krcmar M. Predicting media use in very young children: the role of demographics and parent attitudes. Commun Stud 2013 Sep;64(4):374-394. [doi: 10.1080/10510974.2013.770408]
31. Lauricella AR, Wartella E, Rideout VJ. Young children's screen time: the complex role of parent and child factors. J Appl Dev Psychol 2015 Jan;36:11-17. [doi: 10.1016/j.appdev.2014.12.001]
32. Coyne SM, Radesky J, Collier KM, et al. Parenting and digital media. Pediatrics 2017 Nov;140(Suppl 2):S112-S116. [doi: 10.1542/peds.2016-1758N] [Medline: 29093044]
33. The state of school education: one year into the COVID pandemic. : Organisation for Economic Co-operation and Development; 2021 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/04/the-state-of-school-education_b929a614/201dde84-en.pdf [accessed 2026-01-09]
34. Bergmann C, Dimitrova N, Alaslani K, et al. Young children's screen time during the first COVID-19 lockdown in 12 countries. Sci Rep 2022 Feb 7;12(1):2015. [doi: 10.1038/s41598-022-05840-5] [Medline: 35132065]
35. Bandura A. Social Learning Theory: Prentice Hall; 1977. URL: https://catalog.nlm.nih.gov/discovery/fulldisplay/alma9915626113406676/01NLM_INST:01NLM_INST [accessed 2026-01-09]
36. Rideout V, Hamel E. The media family: electronic media in the lives of infants, toddlers, preschoolers and their parents. : The Kaiser Family Foundation; 2006 URL: https://www.kff.org/wp-content/uploads/2013/01/7500.pdf [accessed 2026-01-09]
37. Kim HN, Lee SA, Lee EK, Chon JH, Kim SH. Effects of the smart media exposure environment on smart media addiction of young children [Article in Korean]. J Korean Soc Child Care 2014;14(4):127-153 [FREE Full text]
38. Oh J, Park Y. A study on preschoolers' smart media use and parents' perception [Article in Korean]. Korean J Child Care Educ Policy 2019;13(3):3-26 [FREE Full text]
39. Poulain T, Ludwig J, Hiemisch A, Hilbert A, Kiess W. Media use of mothers, media use of children, and parent-child interaction are related to behavioral difficulties and strengths of children. Int J Environ Res Public Health 2019 Nov 22;16(23):4651. [doi: 10.3390/ijerph16234651] [Medline: 31766650]

40. Pequegnat W, Szapocznik J. The role of families in preventing and adapting to HIV/AIDS: issues and answers. In: Working with Families in the Era of HIV/AIDS: Sage; 2000:3-26 URL: https://books.google.co.in/books?id=HvQJAZIL12UC&source=gbs_book_other_versions [accessed 2026-01-09]
41. Rasmussen EC, White SR, King AJ, Holiday S, Densley RL. Predicting parental mediation behaviors: the direct and indirect influence of parents' critical thinking about media and attitudes about parent-child interactions. J Media Lit Educ 2016;8(2):1-21. [doi: 10.23860/JMLE-2016-08-02-01]
42. Arnold LB. Family Communication: Theory and Research: Pearson Education Inc; 2007. URL: https://www.biblio.com/book/family-communication-theory-research-arnold-lorin/d/1449710603?srsltid=AfmBOoqtDQCOnxEkRrRVlGHDcUKRBpbZMqYpkg333S6H3dAxSOQFfnJY [accessed 2026-01-09]
43. American Academy of Pediatrics, Committee on Public Education. American Academy of Pediatrics: children, adolescents, and television. Pediatrics 2001 Feb;107(2):423-426. [doi: 10.1542/peds.107.2.423] [Medline: 11158483]
44. Vandewater EA, Rideout VJ, Wartella EA, Huang X, Lee JH, Shim MS. Digital childhood: electronic media and technology use among infants, toddlers, and preschoolers. Pediatrics 2007 May;119(5):e1006-e1015. [doi: 10.1542/peds.2006-1804] [Medline: 17473074]
45. Chakroff JL, Nathanson AI. Parent and school interventions: mediation and media literacy. In: Calvert SL, Wilson BJ, editors. The Handbook of Children, Media, and Development: Blackwell Publishing; 2008:552-576. [doi: 10.1002/9781444302752]
46. Council on Communications and Media, Brown A. Media use by children younger than 2 years. Pediatrics 2011 Nov;128(5):1040-1045. [doi: 10.1542/peds.2011-1753] [Medline: 22007002]
47. Sanders W, Parent J, Forehand R, Sullivan ADW, Jones DJ. Parental perceptions of technology and technology-focused parenting: associations with youth screen time. J Appl Dev Psychol 2016;44:28-38. [doi: 10.1016/j.appdev.2016.02.005] [Medline: 27795603]
48. Joe SS, Kim BH. A study of factors affecting the amount of children's YouTube use. J Korean Entertain Ind Assoc 2021 Oct 31;15(7):45-57. [doi: 10.21184/jkeia.2021.10.15.7.45]
49. Coyne SM, Rogers A, Holmgren HG, et al. Masters of Media: a longitudinal study of parental media efficacy, media monitoring, and child problematic media use across early childhood in the United States. J Child Media 2023;17(3):318-335. [doi: 10.1080/17482798.2023.2200958] [Medline: 37841526]
50. Huang S, Lai X, Li Y, et al. Does parental media mediation make a difference for adolescents? Evidence from an empirical cohort study of parent-adolescent dyads. Heliyon 2023 Apr;9(4):e14897. [doi: 10.1016/j.heliyon.2023.e14897] [Medline: 37095946]
51. Fam JY, Männikkö N, Juhari R, Kääriäinen M. Is parental mediation negatively associated with problematic media use among children and adolescents? A systematic review and meta-analysis. Can J Behav Sci 2023;55(2):89-99. [doi: 10.1037/cbs0000320]
52. Baumrind D. Current patterns of parental authority. Dev Psychol 1971;4(1, Pt.2):1-103. [doi: 10.1037/h0030372]
53. Maccoby EE, Martin JA. Socialization in the context of the family: parent-child interaction. In: Mussen PH, Hetherington EM, editors. Handbook of Child Psychology: Wiley; 1983, Vol. IV:1-101 URL: https://agris.fao.org/search/en/providers/122535/records/65dde6b74c5aef494fd6d65f [accessed 2026-01-09]
54. Van Leeuwen KG, Vermulst AA. Some psychometric properties of the Ghent Parental Behavior Scale. Eur J Psychol Assess 2004 Jan;20(4):283-298. [doi: 10.1027/1015-5759.20.4.283]
55. Padilla-Walker LM, Coyne SM. "Turn that thing off!" parent and adolescent predictors of proactive media monitoring. J Adolesc 2011 Aug;34(4):705-715. [doi: 10.1016/j.adolescence.2010.09.002] [Medline: 20943262]
56. Wood CT, Skinner AC, Brown JD, et al. Concordance of child and parent reports of children's screen media use. Acad Pediatr 2019 Jul;19(5):529-533. [doi: 10.1016/j.acap.2019.04.002] [Medline: 30981024]
57. Elias N, Sulkin I. YouTube viewers in diapers: an exploration of factors associated with amount of toddlers' online viewing. Cyberpsychol J Psychosoc Res Cyberspace 2017;11(3). [doi: 10.5817/CP2017-3-2]
58. Nikken P, Jansz J. Developing scales to measure parental mediation of young children's internet use. Learn Media Technol 2014 Apr 3;39(2):250-266. [doi: 10.1080/17439884.2013.782038]
59. Paruthi S, Brooks LJ, D'Ambrosio C, et al. Recommended amount of sleep for pediatric populations: a consensus statement of the American Academy of Sleep Medicine. J Clin Sleep Med 2016 Jun 15;12(6):785-786. [doi: 10.5664/jcsm.5866] [Medline: 27250809]
60. Arbuckle JL. 7.0 user's guide. : Amos Development Corporation/SPSS Inc; 2006 URL: http://www.hsc.edu.kw/student/app_manuals/SPSS/Amos%207.0%20User's%20Guide.pdf [accessed 2026-01-09]
61. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
62. Buijzen M, Valkenburg PM. Parental mediation of undesired advertising effects. J Broadcast Electron Media 2005 Jun;49(2):153-165. [doi: 10.1207/s15506878jobem4902_1]
63. Wu CST, Fowler C, Lam WYY, Wong HT, Wong CHM, Yuen Loke A. Parenting approaches and digital technology use of preschool age children in a Chinese community. Ital J Pediatr 2014 May 7;40:44. [doi: 10.1186/1824-7288-40-44] [Medline: 24887105]

64. Rosenman R, Tennekoon V, Hill LG. Measuring bias in self-reported data. Int J Behav Healthc Res 2011 Oct;2(4):320-332. [doi: 10.1504/IJBHR.2011.043414] [Medline: 25383095]
65. Byrne R, Terranova CO, Trost SG. Measurement of screen time among young children aged 0-6 years: a systematic review. Obes Rev 2021 Aug;22(8):e13260. [doi: 10.1111/obr.13260] [Medline: 33960616]
66. Barr R, Kirkorian H, Radesky J, et al. Beyond screen time: a synergistic approach to a more comprehensive assessment of family media exposure during early childhood. Front Psychol 2020;11:1283. [doi: 10.3389/fpsyg.2020.01283] [Medline: 32754078]

Abbreviations

CFI: confirmatory fit index
GFI: goodness-of-fit index
NFI: normed fit index

Edited by H Witteman; submitted 06.Apr.2025; peer-reviewed by G Taylor, R Juhari; revised version received 18.Dec.2025; accepted 22.Dec.2025; published 28.Jan.2026.

Please cite as:
Kim JY, Yang AJ, Lee HE
Parental Influence on Children’s Media Use in South Korea: National Population-Based Study
JMIR Hum Factors 2026;13:e75292
URL: https://humanfactors.jmir.org/2026/1/e75292
doi:10.2196/75292

An Internet-Delivered Intervention to Reduce LGBTQ+ Prejudice Among Romanian Teachers: Randomized Controlled Trial

Nastasia Sălăgean[1,2*], PhD; Ioana Maria Latu[2,3*], PhD; Torill Marie Bogsnes Larsen[4], PhD; Andreea Bogdana Isbă oiu[2,5], PhD; Florin Alin Sava[2], PhD

[1] Department of Scientific Research in Economy, Law and Human-Environment Interaction, Institute for Advanced Environmental Research, West University of Timi oara, Boulevard V Pârvan 4, Timisoara, Romania

[2] Department of Psychology, Faculty of Sociology and Psychology, West University of Timisoara, Timisoara, Romania

[3] School of Psychology, Queen's University Beflast, Beflast, United Kingdom

[4] Department of Health Promotion and Development, Faculty of Psychology, University of Bergen, Bergen, Norway

[5] Department of Psychology and Education Sciences, Transilvania University of Brasov, Brasov, Romania

[*] these authors contributed equally

Corresponding Author:
Nastasia Sălăgean, PhD
Department of Scientific Research in Economy, Law and Human-Environment Interaction, Institute for Advanced Environmental Research, West University of Timi oara, Boulevard V Pârvan 4, Timisoara, Romania

Abstract

Background: Discrimination can greatly impact both physical and mental health due to frequent stressors. Younger individuals, particularly those under the age of 17 years, are more adversely affected by victimization. Within the European Union, Romania exhibits poor rankings concerning LGBTQ+ (lesbian, gay, bisexual, transgender/transsexual, queer, and other minority sexual orientations and gender identities) inclusion, with large numbers of LGBTQ+ teenagers experiencing bullying due to their sexual orientation. Given that much of this discrimination and harassment occurs within schools, teachers and counselors are vital in affecting institutional change.

Objective: This study aims to investigate the impact of an intervention on reducing prejudice against the LGBTQ+ community among Romanian teachers and counselors. Most prior interventions of this nature target Western, educated, industrialized, rich, and democratic populations.

Methods: In this randomized controlled trial, we recruited 175 Romanian teachers via a national closed online user group and assigned them to either the experimental or control condition. Participants in the experimental condition received the intervention first and then completed the web-based outcome measures, while those in the control condition completed the measures first and then received the intervention. The intervention, designed for internet-based delivery, consisted of a 1-hour video session led by a pair of researchers. It blended educational information with testimonials of LGBTQ+ people, perspective-taking tasks, and a self-efficacy exercise. We measured LGBTQ+ prejudice (using Attitudes Toward Lesbians and Gay Men Scale, Homophobia Scale, and Attitudes Toward Homosexuals Scale), behavioral intentions, self-efficacy, perspective taking, intergroup disgust sensitivity, intergroup anxiety, empathy, factual knowledge about LGBTQ+ issues, as well as participants' feelings toward lesbian, gay, and bisexual individuals.

Results: Participants in the experimental group (n=89) showed significant reductions in prejudice when using the Attitudes Toward Lesbians and Gay Men Scale ($F_{1,173}$=7.22; *P*=.008) when compared to the control group (n=86), but not when using the other 2 attitudinal scales. We also found that the experimental group had warmer feelings ($F_{1,173}$=4.40; *P*=.04; *d*=0.32), were more likely to engage in supportive behaviors ($F_{1,173}$=13.96; *P*<.001; *d*=0.56), displayed more self-efficacy ($F_{1,173}$=9.14; *P*=.001; *d*=0.33), had more factual knowledge ($F_{1,173}$=11.98; *P*=.001; *d*=0.52), and had a higher ability to take the LGBTQ+ perspective after controlling for contact ($F_{1,172}$=4.77; *P*=.03; *d*=0.28). We did not observe significant differences in terms of intergroup disgust sensitivity ($F_{1,173}$=0.816; *P*=.37), intergroup anxiety for either positive ($F_{1,173}$=.383; *P*=.54) or negative emotions ($F_{1,173}$=0.51; *P*=.48), or empathy ($F_{1,173}$=0.02; *P*=.89).

Conclusions: The intervention offers initial evidence for the effectiveness of a cost-effective and portable online resource for educators and high school counselors, particularly in regions where negative attitudes toward the LGBTQ+ community are prominent. The results show that integrating blended cognitive (information), affective (indirect contact and perspective taking), and behavioral (self-efficacy and empowerment) approaches is a promising avenue for intervention in producing positive outcomes related to LGBTQ+ issues within the school environment.

Trial Registration: ISRCTN Registry ISRCTN 84290049; https://www.isrctn.com/ISRCTN84290049

International Registered Report Identifier (IRRID): RR2-10.2196/54254

(JMIR Hum Factors 2026;13:e63787) doi:10.2196/63787

KEYWORDS

LGBT; discrimination; minority stress; internet; intervention; randomized controlled trial; Romanian teachers; Romania; school teacher; web-based; internet-based delivery; anxiety; empathy; self-efficacy; educators; online session; school; LGBTQ+

Introduction

Background

The LGBTQ+ (lesbian, gay, bisexual, transgender/transsexual, queer, and other minority sexual orientations and gender identities) community has historically faced great adversity, as centuries of prosecution have only recently begun to be overturned. The collapse of communism in Eastern Europe brought several changes in civil societies after they began shifting their focus to prioritizing human rights values and norms [1]. In Romania, homosexuality was officially decriminalized in 2001 because of external and international pressures [1]. Despite these changes, the consequences of the decades of oppression are seen even today, as attitudes toward LGBTQ+ individuals remain overall negative [2]. In fact, when it comes to LGBTQ+ inclusion, Romania tends to perform poorly compared to other EU countries. According to a recent survey [3,4], 45% of LGBTQ+ individuals in Romania reported experiencing discrimination in at least one aspect of their lives, but only 8% reported their experiences to an organization dealing with such issues. Additionally, 43% of the respondents said they faced harassment due to their sexual orientation or gender identity, the second-highest rate among all the EU countries.

The rates mentioned above are particularly concerning when viewed from the perspective of the minority stress model [5]. This model proposes that minorities face unique stressors because of stigmatization and discrimination, and that these can lead to detrimental effects on health. The stress caused by repeated experiences of discrimination can accumulate over time and negatively impact both physical and mental health in the long run.

Several studies have been conducted to test this model. In terms of physical health, a systematic review of 26 studies [6] found that minority stress impacted several physical health outcomes in sexual minorities, including cancer incidence, changes in cardiovascular function, and immune response. Additionally, a different systematic review [7] found that sexual minorities face significantly elevated risks for cardiovascular diseases. Similarly, mental health issues in LGBTQ+ individuals have been widely studied. A systematic review [8] revealed that sexual minorities had a higher risk for mental health issues, substance abuse, and suicide risk. In line with the minority stress model, evidence suggests that individuals who identify as LGBTQ+ also face several obstacles that negatively impact their mental health, such as discrimination, emotional distress, victimization, and barriers to accessing mental health services [9].

Moreover, although data suggest that minority stress affects all LGBTQ+ individuals, it seems that age impacts the stress felt by sexual minorities. Several studies have shown that younger individuals generally report more stress than older ones [10,11]. A meta-analysis revealed that victimization had more negative effects on LGBTQ+ individuals under the age of 17 years as compared to those aged more than 17 years [12]. These findings are corroborated by the impact on suicide risk. A meta-analysis of 35 studies [13] showed that suicide risk was higher for young LGBTQ+ individuals aged 12-20 years compared to cisgender straight young people. Furthermore, sexual minority youth are more likely to experience mental health issues such as anxiety, depression, and impaired academic performance [14]. This effect is especially prevalent within school environments, as supported by a systematic review showing significant negative effects of bias-based bullying for minorities, including LGBTQ+ students [15].

In fact, 44% of LGBTQ+ pupils aged 15 to 17 years in Romania have admitted to hiding their sexual orientation or gender identity in school, the third highest rate in the European Union after Croatia (51%) and Cyprus (47%), and significantly higher than in countries like the Netherlands (16%), France (20%), or Denmark (20%). Disturbingly, half of the respondents aged 15 to 17 years (50%) said that they were bullied at school because of their sexual orientation [3,4]. This finding is further supported by the United Nations Educational, Scientific and Cultural Organization report on school violence [16]. As such, given that discrimination and harassment against LGBTQ+ individuals are quite prevalent in Romanian schools and among students [3,4], it is therefore imperative to create programs that help reduce the stigma and prejudice that Romanian LGBTQ+ students face.

We decided to focus such efforts on teachers as they are perhaps the most significant group that can affect institutional change. Teachers can influence their students' experiences by creating inclusive classroom norms, making these interventions most effective. In fact, survey evidence suggests that supportive communities and schools buffer the negative effects of bias-based bullying for LGBTQ+ students in schools [17] . This finding is backed up by qualitative data, with interviews in schools showing the importance of teacher attitudes and education in creating safe and inclusive environments for LGBTQ+ students [18]. Overall, a recent rapid realist review of interventions to promote LGBTQ+ inclusivity in schools [19] showed that interventions work best when the staff in the school are trained, including education on sexuality, gender issues, as well as how to be an effective ally. Teacher training was related to LGBTQ+ students experiencing less victimization, increased safety, greater self-esteem, improved mental health, and improved academic performance. Moreover, teacher training improved the effectiveness of teacher involvement in gay-straight alliances, which directly impacts LGBTQ+ youth's outcomes in schools [20].

In this paper, we aim to present the findings of a teacher-oriented intervention whose protocol has already been published [21]. The main objective of the intervention is to use a 1-hour online session that focuses on education and contact as primary training elements. The session is meant to help high school teachers cope with cases of witnessing bullying acts against LGBTQ+ students and to improve their knowledge and attitudes toward LGBTQ+ students.

To create a safe and inclusive learning environment for LGBTQ+ students, it is crucial to not only address teacher biases toward them, but also to equip teachers with the skills and knowledge needed to intervene in LGBTQ+-related victimization. This includes providing teachers with training on how to recognize and respond to instances of LGBTQ+ bullying or discrimination, as well as model appropriate behaviors and attitudes toward LGBTQ+ students. By taking a proactive and comprehensive approach to support LGBTQ+ students, we can help ensure that they feel valued, respected, and included in the classroom, thus reducing the risk of significant mental and physical health costs.

According to a meta-analysis [22], interventions can be effective in reducing sexual prejudice. The effect size of these interventions can range from one-third to one-half of an SD. Furthermore, the analysis highlights the most effective strategies for reducing different outcomes. These strategies include educational interventions, contact with LGBTQ+ individuals, and a combination of education and contact. It is worth noting that the majority of interventions aimed at reducing homophobia have been conducted among undergraduate students in Western countries, with none conducted in Eastern Europe or other regions that still harbor clear animosity and prejudice toward the LGBTQ+ community. While some recent research has tested interventions in other countries such as Jamaica [23] or Brazil [24], in Romania (or Eastern Europe), no such interventions have been tested, particularly on teachers.

However, equally important is how the effect of these interventions is measured. The term "homophobia" has garnered a lot of attention since its inception in the 1970s [25]. Several researchers [26,27] have argued the term should not be used because of its inaccuracy, as the inclusion of the word "phobia" would suggest an anxiety-related measure rather than an attitudinal one. Additionally, the misnomer would mainly focus on attitudes toward gay men and leave out lesbians and bisexuals. This conceptual confusion has seeped into the measurements used when measuring the impact of the interventions. Currently, several measures exist that claim to test sexual prejudice, but few interventions have used more than 1 instrument. It would be, therefore, beneficial to include several measures to see which, if any, would better capture the effects of the intervention. Besides prejudice, behavioral, cognitive, and emotional outcomes are also vital in predicting and ensuring significant changes in the classroom.

In order to ensure the success of interventions, it is crucial to consider the cultural and institutional context in which they are implemented. A recent qualitative analysis [28] has shown that participants frequently criticize interventions for their lack of alignment with the context in which they are conducted. This can be used as a justification for resisting change. To achieve better outcomes, interventions should be tailored to the community's unique characteristics, including their beliefs, values, traditions, and social norms.

Aims and Hypotheses

The main objective of this research is to test an intervention plan that aims to improve Romanian teachers' LGBTQ+ outcomes. Our intention is to consider particular cultural and institutional characteristics, rather than simply implementing previous intervention strategies. Our approach involves using education and contact as primary training elements, consistent with meta-analytic findings [22] on training effectiveness. However, we design the educational components according to the specific needs of our target group. For example, we incorporate information on the biological (rather than social) causes of homosexuality, as this is a common misbelief among Romanians. The protocol for this study is peer-reviewed and published in Research Protocols [21].

Additionally, we addressed potential feelings or perceptions of threat (such as the belief that exposure to LGBTQ+ individuals will "cause" children to become gay) as, according to intergroup threat theory [29], these are important predictors of anti-gay bias. We included several elements that had been found useful in other interventions, such as perspective-taking [30] and self-efficacy [31,32]. We also ensured that these elements were culturally appropriate. The trial protocol details were reported elsewhere [21], including a complete list of intervention components, subcomponents, and contents. Further details are given in Table 1 in a previous research paper [21].

Through a randomized controlled trial, we aim to measure the impact of the intervention on the attitudes, behavior, cognition, and emotions of teachers toward the LGBTQ+ individuals compared to those randomly assigned to the control condition. Additionally, we also measure the participants' factual knowledge about LGBTQ+ issues in the classroom, as well as their attitudes toward LGBTQ+ individuals.

Methods

Ethical Considerations

The ethical review committee of the West University of Timisoara, Romania, approved this study (notice 74505/10.11.2022) after reviewing the procedure, measures, and materials used. Each participant was informed of the data collected and the details of the intervention on 3 separate occasions: upon signing up for the intervention session, upon beginning the intervention, and when responding to the questionnaires. Participants provided informed consent by clicking a radio button on an online sign-up page containing written information about the study.

Upon finishing the online multimedia intervention, respondents received an automatic unique code that they had to submit in a separate form to confirm their full participation in the intervention. Once participants confirmed their full participation, they were sent a certificate of participation that could be used to earn continuing education credits. They were also given the opportunity to enter a raffle to win a gift certificate worth 500 RON (equivalent to US $110 or €100).

All research data (outcome measures) were collected anonymously without recording any personal or identifiable information. However, to ensure that participants were indeed teachers, and to generate the participation certificates, the researchers collected personal information (email and name) that was stored in a separate database and never associated with the actual study results.

Participant Recruitment and Eligibility Criteria

By working with the local center for educational resources and assistance to distribute an online message to all their national members, we aimed to recruit teachers or counselors employed in Romanian schools. The interested teachers and school counselors had the opportunity to sign up for 1 of the 17 advertised sessions across 3 months (from December 2022 to February 2023), scheduled at different times during the day, depending on their availability. To ensure more participation, the teachers who signed up for a session were sent a reminder email 24 hours before the commencement of the session.

All participants had to be either teachers or school counselors and be fluent in Romanian, as the intervention was conducted in this language. Additionally, because the trial was advertised and conducted online, respondents had to have basic computer or internet literacy to access the recruitment form and the Zoom meeting links (Zoom Communications, Inc).

Study Design

The experimental design used in this intervention was a randomized controlled trial. We randomly assigned participants to either the experimental or control condition using a 2-group design. Those in the experimental condition received the intervention first and then completed the outcome measures, while those in the control condition completed the outcome measures first and then received the intervention. All participants had to fill out the outcome measures through the use of online questionnaires. This design was chosen for ethical reasons so that all participants would benefit from the intervention and associated resources by the end of their participation. Given the rigorous random assignment to the training and control conditions, we did not expect any baseline differences to influence the outcome measures. We did, however, conduct analyses with and without controlling for habitual contact with LGBT individuals. We chose this approach because discussions in the contact literature suggested that baseline levels of contact can potentially lead to a selection bias of participants to these studies [33,34] and as such, controlling those levels can help us isolate the intervention's specific effect beyond habitual exposure.

Participants were grouped into sessions of up to 30 individuals, depending on their availability. Cluster randomization was done at the session level by the lead researcher, with each session being randomly assigned to either the experimental or control condition. The randomization was done with the use of an online number generator. Individual randomization within the session was not possible given that the intervention was presented to all participants at the same time.

Sessions were scheduled outside of typical working hours on different days and times to ensure there were no systematic biases due to participants' session choice.

This study did not involve any risk to the participants' physical or mental integrity, and no unintended harms were observed during implementation.

In addition to being published in an academic journal [21], the study protocol was also registered in an international clinical study registry (ISRCTN84290049) [35]. The study is reported in accordance with the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines as per the completed checklist (Checklist 1).

Intervention

To ensure a standardized experience for all participants, regardless of the session they attended, each session was led by a pair of researchers who were tasked with delivering scripted instructions, answering participant queries, and ensuring that the study was completed at a consistent pace with minimal distractions.

The stand-alone intervention was designed for internet-based delivery and was multimedia in nature. It contained a recorded animated presentation that contained information on terms, threat reduction, and effects of stigma, behavioral tools, testimonials of LGBTQ+ people, a perspective-taking task, and a self-efficacy exercise. In total, the intervention lasted 50 minutes. The full details on all the components of the intervention are presented in Table 1 in a previous research paper describing the protocol in detail [21].

Sample Size Estimation

In order to determine the appropriate sample size for our study, we undertook a power analysis. Drawing on an average effect size of d=0.66 from a meta-analysis by Bartos et al [22], we aimed to achieve a statistical power of 0.80 as per the Cohen [36] recommendation that type II error should be limited to a probability of 0.20. Our analysis indicated a required sample size of 122 participants, although we elected to over-recruit in order to account for multiple outcomes. There was no pre-established stopping rule, and recruitment continued until all interested participants had been given the opportunity to participate in 1 of the 17 scheduled sessions.

All participants who completed their participation in the study were included in the data analysis, with no exclusions.

Outcomes

Overview

In order to ensure the reliability of the final scores for each outcome, we verified that the Cronbach α coefficient passed the 0.7 threshold. Additionally, to ensure the accuracy of the results in our research, we used 3 different scales in measuring sexual prejudice and antigay bias.

Attitudes Toward Lesbians and Gay Men Scale: Measure of Sexual Orientation Prejudice

This 10-item scale measured beliefs and attitudes toward gay men and lesbians ("I think male homosexuals are disgusting" and "Sex between two women is just plain wrong.") [37]. Items were rated on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). After reverse coding 4 items, scores were averaged into a final score with higher values denoting more negative attitudes. Reliability for this measure was very good (Cronbach α=0.88).

Homophobia Scale (HS): Attitudes Toward Gay Individuals

This scale was composed of 25 items that assessed attitudes toward gay individuals, as well as social avoidance and aggression toward them [38]. Examples of items from the scale included, "Gay people make me nervous" and "I make derogatory remarks about gay people." Participants rated each item on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). After reverse coding 9 items, the scores were averaged to obtain a final score. Higher scores indicate more negative attitudes toward gay people. The scale has shown excellent reliability with a Cronbach α of 0.93.

Attitudes Toward Homosexuals Scale: Assessment of Negative Beliefs

We used an additional scale to measure the attitudes toward and avoidance of gay people [39]. The 12 items included statements such as "Homosexuality is disgusting in the eyes of God," and "If I can, I prefer to not be in the company of homosexuals." Participants rated the statements on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Out of all, 5 items were reverse scored, and the statements were averaged into a final score with higher values denoting more negative attitudes. Reliability for this measure was excellent (Cronbach α=0.91).

Self-Efficacy

The original 10-item scale [40] was adapted for working with LGBTQ+ students in a school setting. The scale measured self-efficacy in dealing with issues related to LGBTQ+ students. Sample items included "If I try hard, I can solve difficult issues related to LGBTQ+ students" and "I can deal with unexpected situations that arise with LGBTQ+ students." Ratings ranged from 1 ("Not at all true for me") to 4 ("Perfectly true for me"), and higher scores indicating more self-efficacy in dealing with LGBTQ+–related behaviors in the classroom. The scale showed an excellent reliability with a Cronbach α of 0.92.

Behavioral Intentions

We used a 16-item scale to evaluate the propensity of teachers to exhibit supportive professional conduct in the classroom pertaining to LGBTQ+ issues [39,40], including discussing queries regarding sexual orientation with students or having books about gay and lesbian issues in the classroom. Each item was rated on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). A final score was obtained by averaging the responses, with higher scores indicating a greater willingness to engage in LGBTQ+ supportive behaviors in the classroom. The measure showed excellent reliability (Cronbach α=0.95).

Factual Knowledge About LGBTQ+ Issues

We designed a questionnaire consisting of 7 questions to evaluate the level of understanding of the participants on LGBTQ+ topics. The questions were carefully crafted based on the training content and covered a range of topics such as the biological basis of gender and the elevated risk of suicide among LGBTQ+ youth. The items were rated as either true or false, and the final score was calculated by adding up the number of correct responses. Participants with higher scores demonstrated a greater knowledge of and familiarity with LGBTQ+ issues.

Feeling Thermometer

We assessed the teachers' attitudes toward gay, lesbian, and bisexual individuals using a feeling thermometer [41]. We asked them to rate their feelings toward each group using a sliding thermometer scale that ranged from 0 (indicating very negative feelings) to 100 (indicating very positive feelings). We then averaged the ratings for the 3 groups into an overall lesbian, gay, and bisexual (LGB) feelings thermometer. The scale showed excellent reliability (Cronbach α=0.98).

Intergroup Disgust Sensitivity

We used a 7-item scale to evaluate the repulsion toward LGBTQ+ groups [42]. The scale included statements such as "After shaking hands with someone who has a different sexual orientation, even if their hands were clean, I would want to wash my hands." Participants rated items on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). One of the items was reverse-scored. The final score was calculated by averaging all the scores, with higher values indicating more disgust toward the LGBTQ+ community. Reliability for this measure was very good (Cronbach α=0.83).

Intergroup Anxiety

This scale consisted of 10 items that were designed to measure anxiety when interacting with people of another sexual orientation [43]. Participants were asked to rate the likelihood of feeling several emotions such as embarrassed, unsure, happy, or accepted, on a scale of 1 (not at all) to 10 (extremely). We calculated 2 scores, one for positive emotions and another for negative emotions. The reliability for both positive (Cronbach α=0.82) and negative emotions (Cronbach α=0.88) the reliability of the scale was very good.

Perspective Taking

This measure consisted of 5 items that assessed how well individuals can empathize with LGBTQ+ people [44]. Participants were asked to rate their ability to understand the issues that are important to LGBTQ+ people. The overall reliability of the scale was very good, with a Cronbach α score of 0.88.

Toronto Empathy Questionnaire

This scale consisted of 16 items that evaluated participants' level of empathy toward others [45]. Examples of statements in the scale included “I feel upset when someone is treated disrespectfully.” Participants were asked to rate the statements on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Among all, 7 items were reverse coded, and the final score was computed by averaging all items. Higher scores indicated higher levels of empathy. Reliability for this measure was very good (Cronbach α=0.80).

Control Variables

Demographics

We asked respondents to indicate their age, gender, and sexual orientation. As all respondents were either teachers or counselors, they are required to have at least a bachelor’s degree in their respective field. Therefore, respondents were not asked to provide their education level.

Contact With LGBT People

We inquired about the frequency of contact (eg, speaking) with individuals who identify as gay, lesbian, bisexual, and transgender on a scale of 1 (almost daily) to 6 (never).

The Duke University Religion Index was used to measure religiosity [46]. The scale included 5 items that assessed religious involvement, organizational and nonorganizational religious activity, as well as subjective religiosity or intrinsic religiosity. The responses to each item were averaged to obtain a single score, which indicated the level of religiosity. Higher scores suggested greater religiosity. Cronbach α for this scale was 0.81, suggesting a very good reliability.

We also asked participants to indicate their political ideology on a 100-point scale from very conservative to very liberal or progressive.

Statistical Analysis

Quantitative data from the surveys were analyzed using SPSS (version 24; IBM Corp) [47]. To examine our hypotheses, a series of one-way between-participants ANOVAs were performed. The objective of this analysis was to compare outcomes between the participants in the experimental and control groups. We also conducted the analyses while controlling for contact with LGBTQ+ individuals (ANCOVA), given that people’s different experiences may influence their response to training, particularly when it comes to affective outcomes. The results were presented as *F* test and *P* values, along with descriptive statistics such as n, mean, and SD. In order to estimate the effect size using means and SDs, we calculated Cohen *d*.

There was no deviation in the analysis from the registered protocol [21]. However, we decided to include a post hoc analysis in Multimedia Appendix 1, to explain some inconsistent results in the case of one outcome—the HS. We also controlled the analyses using the Benjamini-Hochberg false discovery rate (FDR) method to correct for any type I errors that could occur when testing for multiple hypotheses [48]. These analyses were conducted using the R programming language (version 4.5.1; R Foundation for Statistical Computing) [49]. The results are presented with both the corrected (noted as P_{FDR}) and uncorrected *P* values.

Results

Descriptive Statistics

A total of 382 individuals signed up to attend one of the advertised sessions. Further examination revealed that 6 of the submissions were duplicated, rendering 376 unique individuals that signed up. Of these, 175 out of 382 (45.81%) participants actually participated and finished the intervention, and 201 individuals did not attend despite repeated reminders. A CONSORT (Consolidated Standards of Reporting Trials) flowchart of the phases of the intervention is provided in Figure 1. A full summary of the descriptive statistics, for the full sample (N=175), as well as for the experimental group (n=89) and control group (n=86), is presented in Table 1. Also included in the table are the statistical analyses made to check whether there were any significant differences between the 2 groups. As can be seen, no significant differences were found between the experimental group and the control group in terms of age, gender, sexual orientation, religiosity, ideology, or LGBT contact. This confirms successful randomization across the 2 experimental conditions.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flowchart.

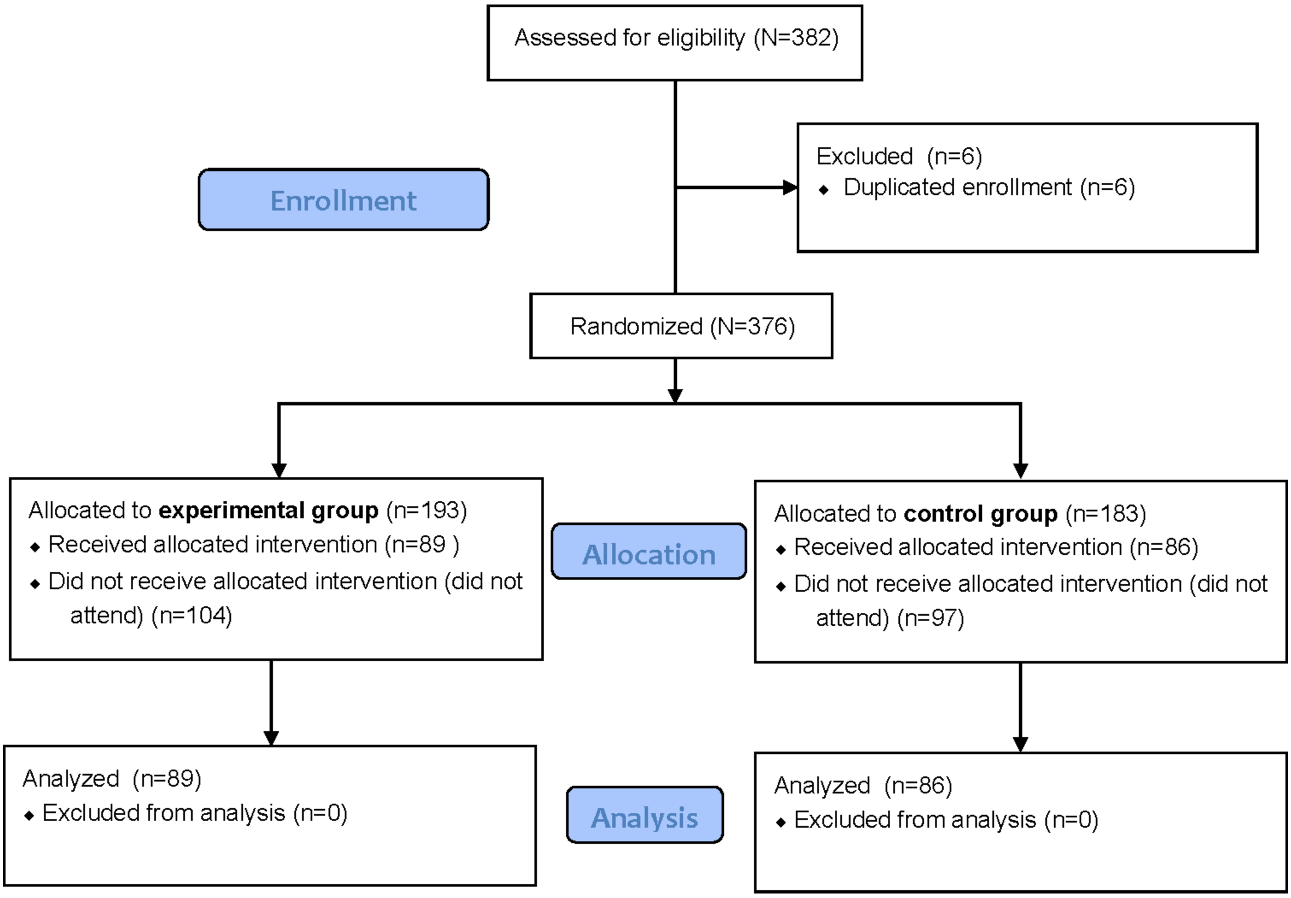


Table . Summary of descriptive statistics (N=175).

Variables	Full sample	Experimental group (n=89)	Control group (n=86)	Chi-square or *t* test (df)	*P* value
Age (y), mean (SD)	43.24 (7.77)	43.34 (7.97)	43.14 (7.62)	0.168 (173)[a]	.87
Gender, n (%)				1.65 (2)[b]	.44
Man	13 (7.4)	8 (9)	5 (5.8)		
Woman	161 (92)	81 (91)	80 (93)		
Nonbinary	1 (0.6)	0 (0)	1 (1.2)		
Sexual orientation, n (%)				4.06 (4)[b]	.40
Heterosexual	146 (83.4)	70 (78.7)	76 (88.4)		
Bisexual	5 (2.9)	3 (3.4)	2 (2.3)		
Other	6 (3.4)	4 (4.5)	2 (2.3)		
Do not know	2 (1.1)	2 (2.2)	0 (0)		
PNTS[c]	16 (9.1)	10 (11.2)	6 (7)		
Religiosity, mean (SD)	3.43 (0.98)	3.46 (0.94)	3.40 (1.03)	0.41 (173)[a]	.69
Ideology, mean (SD)	62.94 (24.86)	65.18 (23.9)	60.62 (25.75)	1.22 (173)[a]	.23
LGBT[d] contact, mean (SD)	5.36 (1.03)	5.39 (0.99)	5.33 (1.08)	0.360 (173)[a]	.72

[a] 2-tailed *t* test (*df*).

[b] Chi-square (*df*).

[c] PNTS: prefer not to say.

[d] LGBT: lesbian, gay, bisexual, and transgender/transsexual.

Outcomes

Significant differences were observed for the Attitudes Toward Lesbians and Gay Men Scale ($F_{1,173}$=7.22; P=.008; P_{FDR}=.02), where teachers in the control group had more negative attitudes (mean 2.70, SD 0.93) than those in the experimental group (mean 2.36, SD 0.77). The effect size, as indicated by Cohen *d*, was small to medium (d=0.41). Teachers who completed the intervention reported less negative attitudes toward lesbians and gays compared to teachers who did not complete the intervention. When controlling for the effect of contact, the experimental condition remained significant ($F_{1,172}$=9.16; P=.003; P_{FDR}=.009; d=0.46).

Interestingly, for the HS, while we did see lower scores for the experimental group (mean 1.95, SD 0.64) than the control group (mean 2.08, SD 0.76), these differences were not statistically significant ($F_{1,173}$=1.58; P=.21; P_{FDR}=.32). This effect remained nonsignificant when controlling for contact ($F_{1,172}$=2.36; P=.13; P_{FDR}=.20). The same can be observed for the Attitudes Toward Homosexuals Scale, where no statistically significant differences ($F_{1,173}$=1.67; P=.20; P_{FDR}=.32) were observed between the experimental group (mean 2.14, SD 0.65) and the control group (mean 2.31, SD 0.88). This effect also remained nonsignificant when controlling for contact ($F_{1,172}$=2.40; P=.12; P_{FDR}=.20).

Initially, significant differences were seen for the LGB feelings thermometer, in that respondents in the experimental group had higher (therefore warmer) feelings toward LGB community members (mean 71.80, SD 27.35) when compared to respondents in the control group (mean 62.64, SD 30.35; $F_{1,173}$=4.40; P=.04; P_{FDR}=.09; d=0.32), though the effect size was small, but after applying the LGB FDR correction, the main result failed to reach significance. When controlling for the effect of contact, the experimental condition remained significant even after applying the FDR correction ($F_{1,172}$=5.82; P=.02; P_{FDR}=.048).

We did not find significant differences between the 2 groups in terms of intergroup disgust sensitivity ($F_{1,173}$=0.82; P=.37; P_{FDR}=.49), intergroup anxiety for either positive ($F_{1,173}$=0.38; P=.54; P_{FDR}=.59) or negative emotions ($F_{1,173}$=0.42; P=.52; P_{FDR}=.59), or empathy ($F_{1,173}$=0.02; P=.89; P_{FDR}=.89). Even when controlling for the effects of contact with LGBTQ+ individuals, we did not see any significant results for intergroup disgust sensitivity ($F_{1,172}$=1.11; P=.29; P_{FDR}=.39), intergroup anxiety for either positive ($F_{1,172}$=0.500; P=.48; P_{FDR}=.52) or negative emotions ($F_{1,172}$=0.51; P=.48; P_{FDR}=.52), or empathy ($F_{1,172}$=0.04; P=.84; P_{FDR}=.84).

However, we did see significant differences in terms of behavioral intentions, in that respondents in the experimental group were more willing to engage in helping behaviors (mean 4.19, SD 0.68) than those in the control group (mean 3.75, SD 0.89; $F_{1,173}$=13.96; P<.001; P_{FDR}=.006; d=0.56), even after controlling for contact ($F_{1,172}$=19.54; P<.001; P_{FDR}=.004).

We also saw that factual knowledge was significantly higher in respondents from the experimental group (mean 5.04, SD 1.24) than those in the control group (mean 4.37, SD 1.32; $F_{1,173}$=11.98; P=.001; P_{FDR}=.006; d=0.52), even when controlling for contact ($F_{1,172}$=12.03; P=.001; P_{FDR}=.004). Furthermore, respondents in the experimental group (mean 3.18, SD 0.42) had significantly higher levels of self-efficacy than those in the control group (mean 2.94; SD 0.93; $F_{1,173}$=9.14; P=.003; P_{FDR}=.01; d=0.33), even after looking at the effects of contact ($F_{1,172}$=11.82; P=.001; P_{FDR}=.004).

Finally, the difference between the experimental and control conditions in terms of perspective taking was just shy of reaching significance ($F_{1,173}$=3.27; P=.07; P_{FDR}=.14; d=0.28), such that respondents in the experimental group had a slightly higher ability to take the LGBTQ+ perspective (mean 3.87, SD 0.69) than those in the control group (mean 3.66, SD 0.80). This analysis only reached significance after controlling for contact ($F_{1,172}$=4.77; P=.03; P_{FDR}=.06), although we note that the effect was small, and after applying the FDR correction, the results were again insignificant.

We conducted the above analyses in line with the preregistered research protocol published in a study by Latu et al [21].

Discussion

Principal Results

Our intervention suggests that even a 1-hour online session that combines informative materials (education, factual knowledge, and behavioral tools to address bullying and support LGBTQ+ students) and testimonials (vicarious contact) produces small to average positive effects on various LGBTQ+ outcomes. The intervention decreased teachers' negative attitude toward LGBTQ+ topics, increased their self-efficacy in dealing with LGBTQ+ issues in the classroom, and led to a higher level of behavioral intentions to engage in supportive behaviors for LGBTQ+ students facing bullying in school contexts. It should be noted that the cognitive and behavioral outcomes were most stable, with the attitudinal outcome being relatively inconsistent across measures and types of analysis. Our inconsistent results across several scales of homophobia or antigay attitudes, together with our supplementary analyses, suggest that latent variables may be at play and that future research should systematically investigate the attitudinal components of homophobia or antigay bias for appropriate measurement.

The online intervention, however, did not lead to significant differences between the experimental and the control groups in terms of intergroup disgust or intergroup anxiety, nor on more general and nonspecific measures such as their empathy level toward other people. This may be because such emotional responses, compared to behavioral and cognitive components, are less malleable especially in the context of LGBTQ+ biases which are deeply ingrained in the Romanian mainstream culture. Our findings may also suggest that it was the educational and behavioral components of our intervention that were most effective. Improved affective outcomes, on the other hand, may result from the contact component of interventions [50]. Although vicarious contact was used in our intervention via recorded video testimonials, more direct and prolonged forms of contact with LGBTQ+ individuals may have the power to change affective responses.

The small to average effect sizes obtained for various outcomes align with an existing meta-analysis [22], in which interventions meant to reduce sexual prejudices obtained similar effect sizes, namely from one-third to one-half of an SD. As already explained elsewhere [21], most of these interventions have been implemented on undergraduate students from Western countries, and none has been implemented in Eastern Europe.

Likewise, our results suggest the opportunity to implement a multipurpose intervention, as suggested in prior studies [24]. Despite its shortness, the intervention showed multiple benefits at the attitudinal, cognitive, affective, and behavioral intention levels.

Taking together, our results are encouraging and a promising starting point for addressing LGBTQ+ biases in teachers. The findings show that an online intervention that is relatively low-cost in terms of personnel and resources could be implemented at a larger scale. Likewise, the session could represent a valuable resource to be added to a more complex intervention. The resource we developed could be an excellent tool for teachers in their professional development training when it comes to topics such as accepting sexual and gender diversity or when required to tackle LGBTQ+–related bullying in school contexts.

Another strength of this work is that the intervention was implemented in a country (Romania) which maintains strong negative attitudes toward LGBTQ+ individuals in the public sphere, with roots in conservative and religious beliefs. Such an educational resource turns away the attention from the moral or religious to psychological aspects such as perspective-taking and diversity acceptance, including vicarious contact with LGBTQ+ people and the problems they face in school settings. The emphasis on the humanistic view in teachers, backed up by education, seems to work even in less tolerant environments concerning LGBTQ+ issues. Therefore, the program behind the current intervention can be seen as a useful and more portable resource for guidance counselors in high schools as well as for other specialists who are interested in upscaling the positive results obtained in various high schools. When evaluating our findings, we believe that the intervention is especially useful given that it led teachers to perceive themselves as more equipped to intervene when LGBTQ+ students experience bullying acts.

Limitations

A significant limitation of this study is the absence of long-term follow-up measurement of outcomes, with measurement being conducted in the same session as the intervention. Although these findings do not allow for inferences about long-term effects, we know that the outcomes of such interventions generally tend to be short-term. For example, a contact intervention [50] showed reduced LGBTQ+ negativity in the short term, but levels returned to baseline 7 days after. In contrast, those in the control condition who were not exposed to the intervention showed even more increased LGBTQ+ negativity 7 days later. Such findings suggest that our intervention, although short-term, can potentially lead to longer-term prevention of worsening of biases. We also suggest that, to secure long-term effects and change inclusivity norms in schools, our intervention could be part of a long-term curriculum of development for teachers.

Another limitation is that there may have occurred biases related to self-selection. Participants were included in the trial based on their voluntary consent and without any certain financial benefit (except for their inclusion in a raffle with a small chance of winning some financial benefits). Although their effort was minimal (attending an online 1-h session excluding the time required for the completion of the outcome questionnaires), it still could have led to the selection of a biased sample from the teachers' population, as most likely teachers with strong negative attitudes toward LGBTQ+ people were more reluctant to attend our study. However, due to random assignment to conditions and the presence of significant effects, we believe there is sufficient evidence about the efficacy of the training.

Another limitation of the study is related to the lack of blinding to the experimental conditions. Because of the nature of the intervention and how it was organized, we could not effectively blind respondents and experimenters to the nature of how the trial worked, thus increasing the risk of a type I error. Future intervention could use artificial intelligence or other automated methods to deliver the intervention in a standardized way, to avoid any demand characteristics or experimenter effects.

A final issue is around the gender composition of our sample, which predominantly consists of women. This gender composition is in line with worldwide trends [51] but also national trends in the Romanian education system [52]. However, theoretically, future studies could investigate the effects of gender on the efficiency of such training tools and adapt them accordingly.

Conclusions

This intervention is a promising resource that can serve as a valuable and relatively low-cost resource (no specialized human intervention is needed) for teachers and high school counselors, particularly in countries where negative attitudes toward LGBTQ+ people prevail. Our study also offers suggestions as to how such interventions could be improved and embedded in a longer-term program to increase classroom inclusiveness for LGBTQ+ students.

Acknowledgments

The authors thank Sebastian E Barto , Lavinia Chirică, and Andre Rădulescu for their advice and assistance in developing the intervention.

Funding

The research leading to these results has recieved funding from the NO grants 2014-2021, under project contract 17/2020.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional analyses.
[DOCX File, 54 KB - humanfactors_v13i1e63787_app1.docx]

Checklist 1
CONSORT-EHEALTH (V 1.6.1) checklist.
[PDF File, 1041 KB - humanfactors_v13i1e63787_app2.pdf]

References

1. Margarit D. LGBTQ rights, conservative backlash and the constitutional definition of marriage in Romania. Gender, Place & Culture 2019 Nov 2;26(11):1570-1587. [doi: 10.1080/0966369X.2019.1567472]
2. Wilson K. Attitudes toward LGBT people and their rights in Europe. In: Oxford Research Encyclopedia: Politics: Oxford University Press; 2020. [doi: 10.1093/acrefore/9780190228637.013.1335]
3. A long way to go for LGBTI equality. European Union Agency for Fundamental Rights. 2020. URL: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-lgbti-equality-1_en.pdf [accessed 2025-07-30]
4. A long way to go for LGBTI equality EU-LGBTI II: country data—Romania. : European Union Agency for Fundamental Rights; 2020 URL: https://fra.europa.eu/sites/default/files/fra_uploads/lgbti-survey-country-data_romania.pdf [accessed 2025-07-30]
5. 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. [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
6. Flentje A, Heck NC, Brennan JM, Meyer IH. The relationship between minority stress and biological outcomes: a systematic review. J Behav Med 2020 Oct;43(5):673-694. [doi: 10.1007/s10865-019-00120-6] [Medline: 31863268]
7. Caceres BA, Brody A, Luscombe RE, et al. A systematic review of cardiovascular disease in sexual minorities. Am J Public Health 2017 Apr;107(4):e13-e21. [doi: 10.2105/AJPH.2016.303630] [Medline: 28207331]
8. Plöderl M, Tremblay P. Mental health of sexual minorities. a systematic review. Int Rev Psychiatry 2015;27(5):367-385. [doi: 10.3109/09540261.2015.1083949] [Medline: 26552495]
9. Moagi MM, van Der Wath AE, Jiyane PM, Rikhotso RS. Mental health challenges of lesbian, gay, bisexual and transgender people: an integrated literature review. Health SA 2021;26(1):1487. [doi: 10.4102/hsag.v26i0.1487] [Medline: 33604059]
10. Vale M, Bisconti T. Risk and resilience: highlighting minority stress in a lifespan sample of sexual minorities. Innov Aging 2022 Dec 20;6(Supplement_1):386-387. [doi: 10.1093/geroni/igac059.1523]
11. Vale MT, Bisconti TL. Age differences in sexual minority stress and the importance of friendship in later life. Clin Gerontol 2021;44(3):235-248. [doi: 10.1080/07317115.2020.1836107] [Medline: 33143546]
12. Fedewa AL, Ahn S. The effects of bullying and peer victimization on sexual-minority and heterosexual youths: a quantitative meta-analysis of the literature. J GLBT Fam Stud 2011 Jul;7(4):398-418. [doi: 10.1080/1550428X.2011.592968]
13. di Giacomo E, Krausz M, Colmegna F, Aspesi F, Clerici M. Estimating the risk of attempted suicide among sexual minority youths: a systematic review and meta-analysis. JAMA Pediatr 2018 Dec 1;172(12):1145-1152. [doi: 10.1001/jamapediatrics.2018.2731] [Medline: 30304350]
14. Oswalt SB, Wyatt TJ. Sexual orientation and differences in mental health, stress, and academic performance in a national sample of U.S. college students. J Homosex 2011;58(9):1255-1280. [doi: 10.1080/00918369.2011.605738] [Medline: 21957858]
15. Sapouna M, de Amicis L, Vezzali L. Bullying victimization due to racial, ethnic, citizenship and/or religious status: a systematic review. Adolescent Res Rev 2023 Sep;8(3):261-296. [doi: 10.1007/s40894-022-00197-2]
16. Behind the numbers: ending school violence and bullying. : United Nations Educational, Scientific and Cultural Organization; 2019 URL: https://www.unicef.org/media/66496/file/behind-the-numbers.pdf [accessed 2025-07-30]
17. Gower AL, Watson RJ, Erickson DJ, Saewyc EM, Eisenberg ME. LGBQ youth’s experiences of general and bias-based bullying victimization: the buffering role of supportive school and community environments. Int Journal of Bullying Prevention 2021 Jun;3(2):91-101. [doi: 10.1007/s42380-020-00065-4] [Medline: 34926991]
18. Steck AK, Perry D. Challenging heteronormativity: creating a safe and inclusive environment for LGBTQ students. J Sch Violence 2018 Apr 3;17(2):227-243. [doi: 10.1080/15388220.2017.1308255]

19. Schlief M, Stefanidou T, Wright T, Levy G, Pitman A, Lewis G. A rapid realist review of universal interventions to promote inclusivity and acceptance of diverse sexual and gender identities in schools. Nat Hum Behav 2023 Apr;7(4):556-567. [doi: 10.1038/s41562-023-01521-z] [Medline: 36781989]
20. Day JK, Fish JN, Grossman AH, Russell ST. Gay-straight alliances, inclusive policy, and school climate: LGBTQ youths' experiences of social support and bullying. J Res Adolesc 2020 Feb;30 Suppl 2(Suppl 2):418-430. [doi: 10.1111/jora.12487] [Medline: 30861243]
21. Latu I, Sălăgean N, Larsen TMB, Isbasoiu AB, Sava FA. Testing the effectiveness of an intervention to improve Romanian teachers' LGBT+-related attitudes, cognitions, behaviors, and affect: protocol for a randomized controlled trial. JMIR Res Protoc 2024 Apr 23;13:e54254. [doi: 10.2196/54254] [Medline: 38652533]
22. Bartoş SE, Berger I, Hegarty P. Interventions to reduce sexual prejudice: a study-space analysis and meta-analytic review. J Sex Res 2014;51(4):363-382. [doi: 10.1080/00224499.2013.871625] [Medline: 24754358]
23. West K. Understanding and reducing sexual prejudice in Jamaica: theoretical and practical insights from a severely anti-gay society. J Sex Res 2018;55(4-5):472-485. [doi: 10.1080/00224499.2017.1416055] [Medline: 29411998]
24. Costa AB, Pase PF, de Camargo ES, et al. Effectiveness of a multidimensional web-based intervention program to change Brazilian health practitioners' attitudes toward the lesbian, gay, bisexual and transgender population. J Health Psychol 2016 Mar;21(3):356-368. [doi: 10.1177/1359105316628748] [Medline: 26987830]
25. Smith KT. Homophobia: a tentative personality profile. Psychol Rep 1971 Dec;29(3):1091-1094. [doi: 10.2466/pr0.1971.29.3f.1091] [Medline: 5139344]
26. Herek GM. Beyond "Homophobia": Thinking about sexual prejudice and stigma in the twenty-first century. Sex Res Soc Policy 2004 Apr;1(2):6-24. [doi: 10.1525/srsp.2004.1.2.6]
27. Schiffman JB, Delucia-Waack JL, Gerrity DA. An examination of the construct of homophobia. J LGBT Issues Couns 2006 May 10;1(1):75-93. [doi: 10.1300/J462v01n01_06]
28. Barto SE, Hegarty P. Negotiating theory when doing practice: a systematic review of qualitative research on interventions to reduce homophobia. J Homosex 2019;66(9):1262-1286. [doi: 10.1080/00918369.2018.1500780] [Medline: 30095380]
29. Stephan WG, Renfro CL, Davis M. The role of threat in intergroup relations. In: Wagner U, Tropp LR, Finchilescu G, Tredoux C, editors. Improving Intergroup Relations: Building on the Legacy of Thomas F Pettigrew: Blackwell Publishing; 2008:55-72. [doi: 10.1002/9781444303117]
30. Tompkins TL, Shields CN, Hillman KM, White K. Reducing stigma toward the transgender community: an evaluation of a humanizing and perspective-taking intervention. Psychol Sex Orientat Gend Divers 2015;2(1):34-42. [doi: 10.1037/sgd0000088]
31. van Zomeren M, Postmes T, Spears R. Toward an integrative social identity model of collective action: a quantitative research synthesis of three socio-psychological perspectives. Psychol Bull 2008 Jul;134(4):504-535. [doi: 10.1037/0033-2909.134.4.504] [Medline: 18605818]
32. Stewart TL, Latu IM, Branscombe NR, Denney HT. Yes we can!: prejudice reduction through seeing (inequality) and believing (in social change). Psychol Sci 2010 Nov;21(11):1557-1562. [doi: 10.1177/0956797610385354] [Medline: 20889931]
33. Pettigrew TF, Tropp LR. A meta-analytic test of intergroup contact theory. J Pers Soc Psychol 2006 May;90(5):751-783. [doi: 10.1037/0022-3514.90.5.751] [Medline: 16737372]
34. Al Ramiah A, Hewstone M. Intergroup contact as a tool for reducing, resolving, and preventing intergroup conflict: evidence, limitations, and potential. Am Psychol 2013 Oct;68(7):527-542. [doi: 10.1037/a0032603] [Medline: 24128316]
35. Testing the effectiveness of an intervention to improve Romanian teachers' outcomes related to LGBT+ inclusion. ISRCTN. 2023. URL: https://www.isrctn.com/ISRCTN84290049 [accessed 2025-12-12]
36. Cohen J. The earth is round (p <.05). Am Psychol 1994;49(12):997-1003. [doi: 10.1037/0003-066X.49.12.997] [Medline: 8122813]
37. Herek GM, McLemore KA. Attitudes Toward Lesbians and Gay Men Scale. In: Davis CM, Yarber WL, Bauserman R, Schreer G, Davis SL, editors. Handbook of Sexuality-Related Measures, 2nd edition: Sage Publications; 1998:392-394.
38. Wright LW, Adams HE, Bernat J. Development and validation of the homophobia scale. J Psychopathol Behav Assess 1999 Dec;21(4):337-347. [doi: 10.1023/A:1022172816258]
39. Altemeyer B. Enemies of Freedom: Understanding Right-Wing Authoritarianism: Jossey-Bass; 1988. URL: https://cmc.marmot.org/Record/.b12238223 [accessed 2025-12-12]
40. Schwarzer R, Jerusalem M. Self-efficacy as a resource factor in stress appraisal processes. In: Schwarzer R, editor. Self-Efficacy: Thought Control of Action: Hemisphere Publishing Corp; 1992:195-213 URL: https://psycnet.apa.org/record/1992-97719-009 [accessed 2025-12-12]
41. Lavrakas PJ. Feeling thermometer. In: Encyclopedia of Survey Research Methods: Sage; 2008:275-277. [doi: 10.4135/9781412963947]
42. Hodson G, Choma BL, Boisvert J, Hafer CL, MacInnis CC, Costello K. The role of intergroup disgust in predicting negative outgroup evaluations. J Exp Soc Psychol 2013 Mar;49(2):195-205. [doi: 10.1016/j.jesp.2012.11.002]
43. Stephan WG, Stephan CW. Intergroup anxiety. J Soc Issues 1985 Oct;41(3):157-175. [doi: 10.1111/j.1540-4560.1985.tb01134.x]

44. Dessel AB. Measuring the effects of intergroup dialogue on teachers' attitudes, feelings, and behaviors regarding lesbian, gay, and bisexual students and parents [Dissertation]. : University of Tennessee; 2008 URL: https://trace.tennessee.edu/utk_graddiss/426/ [accessed 2025-12-12]
45. Spreng RN, McKinnon MC, Mar RA, Levine B. The Toronto Empathy Questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures. J Pers Assess 2009 Jan;91(1):62-71. [doi: 10.1080/00223890802484381] [Medline: 19085285]
46. Koenig HG, Büssing A. The Duke University Religion Index (DUREL): a five-item measure for use in epidemological studies. Religions (Basel) 2010;1(1):78-85. [doi: 10.3390/rel1010078]
47. IBM SPSS Statistics for Windows, Version 240. IBM Corp. 2016. URL: https://www.ibm.com/support/pages/release-notes-ibm-spss-statistics-240 [accessed 2025-12-12]
48. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Statist 2001;29(4):1165-1188. [doi: 10.1214/aos/1013699998]
49. R Core Team. R: a language and environment for statistical computing. : R Foundation for Statistical Computing; 2025 URL: https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf [accessed 2025-12-12]
50. Cramwinckel FM, Scheepers DT, Wilderjans TF, de Rooij RJB. Assessing the effects of a real-life contact intervention on prejudice toward LGBT people. Arch Sex Behav 2021 Oct;50(7):3035-3051. [doi: 10.1007/s10508-021-02046-0] [Medline: 34505215]
51. Global education monitoring report 2020: gender report, a new generation: 25 years of efforts for gender equality in education. : United Nations Educational, Scientific and Cultural Organization; 2020 URL: https://www.unesco.org/gem-report/en/publication/new-generation-25-years-efforts-gender-equality-education [accessed 2025-10-09]
52. Report on the state of pre-university education in Romania 2019 [Report in Romanian]. : Ministerul Educatiei; 2020 URL: https://www.edu.ro/sites/default/files/_fi%C8%99iere/Minister/2020/Transparenta/Stare%20invatamant/Stare%20preuniversitar_rev_5.07.2021.pdf [accessed 2025-10-09]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
FDR: false discovery rate
HS: Homophobia Scale
LGB: lesbian, gay, and bisexual
LGBTQ+: lesbian, gay, bisexual, transgender/transsexual, queer, and other minority sexual orientations and gender identities

Edited by A Kushniruk; submitted 29.Jun.2024; peer-reviewed by R Willems; revised version received 03.Oct.2025; accepted 20.Nov.2025; published 16.Jan.2026.

Please cite as:
Sălăgean N, Latu IM, Larsen TMB, Isbă oiu AB, Sava FA
An Internet-Delivered Intervention to Reduce LGBTQ+ Prejudice Among Romanian Teachers: Randomized Controlled Trial
JMIR Hum Factors 2026;13:e63787
URL: https://humanfactors.jmir.org/2026/1/e63787
doi:10.2196/63787

Demystifying Quality Metrics and Unveiling the True Measure of Quality of Care in Nursing Homes: Mixed Effects Analysis

Suhas Bharadwaj[1,2*], PhD; Haneen Ali[3*], PhD

[1]Department of Industrial and Systems Engineering, College of Engineering, Auburn University, Auburn, AL, United States

[2]Department of Industrial and Systems Engineering, College of Science and Engineering, University of Minnesota, Twin Cities, Minneapolis, MN, United States

[3]Mechanical and Industrial Engineering Department, Faculty of Engineering and Technology, Applied Science Private University, 21 Al Arab Street, Amman, Jordan

[*]all authors contributed equally

Corresponding Author:
Haneen Ali, PhD
Mechanical and Industrial Engineering Department, Faculty of Engineering and Technology, Applied Science Private University, 21 Al Arab Street, Amman, Jordan

Abstract

Background: The 5-Star Quality Rating System for nursing homes plays a central role in evaluating quality of care, although it has both strengths and limitations. This system relies heavily on the Minimum Data Set and derives several quality measures (QMs) from it. In this study, we validated the effectiveness of the 5-Star Quality Rating System for nursing homes and its underlying QMs in estimating quality of care. We constructed a panel dataset of US nursing homes (n=15,416) active from May 2020 to June 2023, retrieving data from three major sources: (1) COVID-19 nursing home data, (2) Payroll-Based Journal data, and (3) nursing home QM snapshots. The outcome variables included (1) resident infection, (2) staff infection, or (3) resident and staff deaths. The predictor variables were the 5-Star Quality Rating System for nursing homes and its underlying QMs classified as structure, process, or outcome (SPO) QMs.

Objective: This study aims to evaluate the effectiveness of nursing home QMs by regressing nursing home COVID-19 outcomes on nursing home QMs, classified using the Donabedian SPO framework. We hypothesized that nursing homes with better structural quality (eg, greater staff availability, better skill mix, and so on), better process quality (eg, lower restraint use and higher vaccination rates), and better outcome quality (eg, lower number of residents with pressure ulcers and a lower number of resident falls) experienced better COVID-19 performance in terms of resident and staff infections and deaths.

Methods: To examine the association between the COVID-19 outcomes and SPO QMs, we imputed missing values in the dataset using random forest. Subsequently, we modeled the imputed dataset using hurdled zero-inflated negative binomial mixed effects models. The zero inflation model included factors influencing initial susceptibility to COVID-19 or factors influencing the possibility of death after COVID-19 had been contracted. The model estimates were conditioned on zero inflation and random effects.

Results: Staffing measures ($P<.001$ for all variables in all models), health deficiency scores ($P<.001$ for all variables in at least 1 model), COVID-19 hospitalization rates ($P<.001$ for all variables in at least 2 models), and vaccinations ($P<.001$ for all variables in at least 2 models) exhibited meaningful relationships with the COVID-19 outcomes, while the 5-star components, Medicaid dependency, and ownership showed no clear relationships.

Conclusions: Although widely used, the 5-Star Quality Rating System for nursing homes is an unreliable performance measure. Concerted efforts from lawmakers, policy makers, and lobbyists are needed to refine and enhance the measure, thereby ensuring its reliability and effectiveness.

(JMIR Hum Factors 2026;13:e72770) doi:10.2196/72770

KEYWORDS

5-Star Quality Rating; nursing homes; performance measures; outcomes; COVID-19; nursing home industry; nursing home performance; nursing home quality; quality measures; SPO; structure, process, or outcome; Donabedian SPO framework

Introduction

In the United States, nursing homes are essential care providers for more than 1.3 million residents across 15,600 facilities [1,2]. As the third-largest sector in the health care industry, nursing homes employ more than 1.7 million dedicated staff members [2] who cater to the complex needs of the older adults, who are often frail and affected by multimorbidity [3-6]. With projections indicating consistent increases in the older adult population [7] and average life expectancy [8,9], the demand for nursing home care is expected to rise sharply [8,10,11]. This makes it crucial for nursing homes to optimize their resource allocation while providing excellent care. Embracing the concept of quality of care (QoC) is not just a strategic necessity for survival in a competitive market but also a moral obligation to ensure the well-being of older adults.

The United States recorded 649,611 excess deaths more than expected in the first year of the COVID-19 pandemic, a 23% increase over the previous year. The largest share of excess deaths occurred among older adults [12]. In the nursing home setting, older adults were disproportionately affected [2] because of a combination of individual factors, such as advanced age [13,14] and comorbidities [14,15], and organizational factors, such as understaffing, crowding [2], limited supply of personal protective equipment [16], and inadequate infection prevention and control readiness [2,17,18]. While the Centers for Disease Control and Prevention proposed rigorous COVID-19 control measures, including maintaining personal hygiene practices, mask use, self-isolation, mobility restrictions, and physical distancing [19,20], they could not compensate for the chronic noncompliance of infection prevention and control practices. A May 2020 United States Government Accountability Office report found that 82% of Centers for Medicare and Medicaid Services (CMS)–certified nursing homes had at least one infection-related deficiency between 2013 and 2017, with nearly half receiving citations over multiple years [17,21].

Aimed at deterring noncompliance and promoting accountability, CMS launched the Nursing Home Care (NHC) website in 2002 to improve transparency by expanding public access to quality information [22]. However, despite CMS's efforts, this website is not well known and is considered difficult to comprehend. When consulting NHC, consumers are unable to discern meaningful differences between nursing homes to make informed decisions about their well-being and that of their loved ones. To address this problem, the CMS introduced the 5-Star Quality Rating System for nursing homes in 2008, summarizing the detailed information provided on NHC [23].

The 5-Star Nursing Homes Quality Rating System for Nursing Homes is based on data from the Minimum Data Set (MDS) and includes several quality measures (QMs) derived from it. Concerns have been raised about the effectiveness of the 5-Star Quality Rating System for nursing homes because of its reliance on the MDS. The MDS has several limitations, including inconsistent reporting due to ambiguous instructions and subjective items [24]. In addition, comparing across facilities proves difficult when the QMs emphasize rare events. Such events usually lead to large SEs and large CIs, making it challenging to assess true quality differences [24]. The QMs are assumed to be linear; however, surprisingly, they are sometimes nonlinear [24,25]. The MDS QMs are also presumed to use the complete spectrum of possible values even in situations that deviate from medical norms, such as pressure ulcer rates below 2% [24,26]. Furthermore, ascertainment bias, a type of detection bias caused by inadequate recording of relevant QMs, and substantial interrater variability often impact the reliability of MDS data [24].

Additionally, nursing home studies face other challenges beyond those associated with the sector's reliance on the MDS. One problem is the absence of resident-level data, which impedes the establishment of causality. Moreover, endogeneity is inadequately controlled due to the nonindependence of dependent variables and error terms, which leads to biased results [27-29]. Additionally, the omission of risk adjustment for relevant confounders may make it impossible to capture any associations between a confounder and the other variables [27,30]. In this study, we used a panel dataset to evaluate the effectiveness of several nursing home QMs in explaining nursing home QoC in terms of COVID-19 health outcomes. We achieved this by regressing COVID-19 outcomes measured in the US nursing homes during the period May 2020 to June 2023 on nursing home QMs, classified using the Donabedian structure-process-outcome (SPO) framework, which are used in constructing the 5-Star Quality Rating System for nursing homes. Our research contributes to the literature in several ways. First, most scholars who used COVID-19 health outcomes have so far considered only a limited number of nursing home QMs. Moreover, they have relied on small sample sizes or used cross-sectional data over longitudinal data. We addressed these limitations by using a panel dataset of COVID-19 data for the nursing homes in the United States collected during the period May 2020 to July 2023. Second, to tackle the challenge of large SEs associated with rare events, we opted for a larger sample size, thus enhancing statistical power and refining our estimates. Third, to mitigate the concern of inadequate control of endogeneity, we used hierarchical generalized linear mixed effects modeling.

We hypothesized that nursing homes with better structural quality (eg, greater staff availability and better skill mix), better process quality (eg, lower restraint use and higher vaccination rates), and better outcome quality (eg, lower number of residents with pressure ulcers and lower number of resident falls) experienced better COVID-19 performance in terms of resident and staff infections and deaths. An empirical analysis of this issue offers fresh perspectives that were previously unavailable from existing research.

Methods

Data Sources

To test the study's hypotheses, we used a panel dataset created using 3 major publicly available data sources: COVID-19 nursing home data, Payroll-Based Journal data, and nursing home QM snapshots.

COVID-19 health outcomes for the period May 2020 to June 2023 were obtained from COVID-19 nursing home data. The latter were obtained as a single dataset containing weekly summaries of resident and staff infections and deaths. Payroll-Based Journal data are collected quarterly and offer daily summaries of employee weekly hours for different staff, including nursing, non-nursing, employee, and contract staff. Nursing home QM snapshots include monthly summaries of QMs obtained from multiple data sources, such as the MDS, claims, penalties, provider information, survey summaries, and vaccination data.

The data were aggregated to create a unified dataset. Data aggregation was performed using 2 unique identifiers: one identified the facility (ie, the federal provider number), whereas the other identified the period (ie, week, month, or quarter). All statistical analyses were performed using R version 4.0.0 or higher (R Foundation for Statistical Computing) [31], and a 2-sided *P* less than .05 was considered statistically significant.

Dependent Variables

The measures “residents weekly confirmed COVID-19,” “residents weekly COVID-19 deaths,” “staff weekly confirmed COVID-19,” and “staff weekly COVID-19 deaths” monitored the number of residents or staff who tested positive for COVID-19 and those who died due to COVID-19 in a particular week. These figures were reported by providers weekly from May 24, 2020, to June 6, 2023, with the data for the week ending May 25, 2020, potentially including reporting from January 1, 2020, to May 24, 2020.

In the analysis, we combined “residents weekly COVID-19 deaths” and “staff weekly COVID-19 deaths” into a single value representing “total (resident and staff) weekly COVID-19 deaths.” We adopted this approach because both variables were characterized by sparsity and a high number of zeros, and combining them in this manner increased statistical power and, subsequently, the variance for modeling of zero-inflated outcome variables [32].

Independent Variables

The set of independent variables was categorized based on the Donabedian SPO framework. According to this framework, the quality of health care, including nursing home care, can be assessed at a facility based on 3 quality components: structure, process, and outcome. *Structure* refers to the settings where health care is provided (eg, buildings and technology infrastructure) and accessibility features. *Process* encompasses the actions taken in giving and receiving care, such as pain management, error prevention, and care follow-ups. *Outcome* indicates the consequences of the provided health care, including mortality rates, readmission rates, and functional status. The quality of health care is contingent upon the interplay of these categories, as Donabedian eloquently stated, “A good structure increases the likelihood of good process, and good process increases the likelihood of good outcomes” [27,33].

Structural QMs

For our analysis, we used a comprehensive array of structural QMs encompassing various nursing home attributes (Textbox 1). To gauge the workforce dynamics, we factored in staffing hours, which were calculated separately for employees and contract staff and for nursing and non-nursing staff; we also examined turnover rates for nursing personnel, registered nurses, and non-nursing staff. Furthermore, we considered characteristics tied to the 5-Star Quality Rating System for nursing homes, including facility fines, penalties, and health and fire safety deficiencies, alongside organizational attributes, such as ownership type, provider category, special focus status, and the presence of resident and family councils, all of which contributed to our assessment of overall performance and stability.

Textbox 1. Structural quality measures (count model predictors). “N” denotes a numeric variable, whereas “C” denotes a categorical variable. The number preceding “C” represents the number of levels in the categorical variable.

Provider information model variables

- Percentage of occupied beds (N)
- Provider type (4C)
- Provider resides in hospital (2C)
- Days since approval to provide Medicare and Medicaid services (N)
- Continuing care retirement community (2C)
- Special focus status (2C)
- Abuse icon (2C)
- Most recent health inspection more than 2 years (2C)
- Provider changed owner in previous 12 months (2C)
- With a resident and family council (4C)

Penalties and staffing model variables

- Employee nursing total weekly hours (N)
- Employee non-nursing total weekly hours (N)
- Contract nursing total weekly hours (N)
- Contract non-nursing total weekly hours (N)

Process QMs

An ample set of process QMs was considered (Textbox 2). Among these were factors pertinent to COVID-19, such as the number of residents hospitalized with confirmed cases and their vaccination status, along with weekly admissions of residents previously treated for the virus. The health care aspects included measures such as the percentage of long-stay residents appropriately receiving pneumococcal and influenza vaccines, the percentage of long-stay residents receiving antianxiety or hypnotic medications, those subjected to physical restraints, and those experiencing adverse effects of in-dwelling catheters. For short-stay residents, the metrics focused on appropriate pneumococcal and influenza vaccination rates and the percentage of residents receiving newly prescribed antipsychotic medications.

Textbox 2. Process quality measures (count model predictors). “N” denotes a numeric variable.

Process model variables

- Residents’ weekly COVID-19 admissions (N)
- Residents hospitalized with confirmed COVID-19 (N)
- Residents hospitalized with confirmed COVID-19 and up to date with vaccines (N)
- Percentage of current residents up to date with COVID-19 vaccines (N)
- Percentage of current health care personnel up to date with COVID-19 vaccines (N)
- Percentage of long-stay residents assessed and appropriately given the pneumococcal vaccine (N)
- Percentage of long-stay residents assessed and appropriately given the seasonal influenza vaccine (N)
- Percentage of long-stay residents who have received an antianxiety or hypnotic medication (N)
- Percentage of long-stay residents who have received an antipsychotic medication (N)
- Percentage of long-stay residents who are physically restrained (N)
- Percentage of long-stay residents with catheters inserted and left in their bladders (N)
- Percentage of short-stay residents assessed and appropriately given the pneumococcal vaccine (N)
- Percentage of short-stay residents who have received a newly prescribed antipsychotic medication (N)
- Percentage of short-stay residents assessed and appropriately given the seasonal influenza vaccine (N)

Outcome QMs

The outcome QMs used for the analysis encompassed various measures of resident care and well-being (Textbox 3). Long-stay measures focusing on depressive symptoms, weight loss, deteriorating mobility, increased need for assistance with daily activities, urinary tract infections (UTIs), and loss of bowel or bladder control were included. Short-stay measures evaluated functional improvements in mobility, outpatient emergency department visits, and rehospitalizations within 30 days of admission.

Textbox 3. Outcome quality measures (count model predictors). “N” denotes a numeric variable.

Outcome model variables

- Number of residents with a new positive COVID-19 test result (N)
- Number of staff and personnel with a new positive COVID-19 test result (N)
- Percentage of SNF residents with new or worsened pressure ulcers (N)
- Percentage of high-risk long-stay residents with pressure ulcers (N)
- Percentage of long-stay residents who have experienced one or more falls with major injury (N)
- Percentage of long-stay residents who have depressive symptoms (N)
- Percentage of long-stay residents who have lost too much weight (N)
- Percentage of long-stay residents whose ability to move independently has worsened (N)
- Percentage of long-stay residents whose need for help with daily activities has increased (N)
- Percentage of long-stay residents with urinary tract infections (N)
- Percentage of low-risk long-stay residents who lose control of their bowels or bladders (N)
- Percentage of short-stay residents who have made improvements in function (N)
- Percentage of short-stay residents who have had an outpatient emergency department visit (N)
- Percentage of short-stay residents who have been rehospitalized after a nursing home admission (N)

Penalties and staffing model variables

- Total fines (N)
- Total amount (N)
- Total penalties (N)
- Total days (N)
- Number of facility-reported incidents (N)
- Total nursing staff turnover (N)
- Registered nurse turnover (N)
- Number of administrators who have left the nursing home (N)

Surveys model variables

- Total health deficiencies inspection cycle 1 (N)
- Total health deficiencies inspection cycle 2 (N)
- Total health deficiencies inspection cycle 3 (N)
- Total fire safety deficiencies inspection cycle 1 (N)
- Total fire safety deficiencies inspection cycle 2 (N)
- Total fire safety deficiencies inspection cycle 3 (N)
- Total weighted health survey score (N)

Data Cleaning

Prior to analysis, we removed all the observations in the dataset that contained a flag for either incorrect data or substandard data. This flagging was part of the Centers for Disease Control and Prevention’s quality assurance check, conducted by CMS on 8 data fields to ensure they did not contain implausible or erroneous values. Our decision to exclude flagged observations mirrors those made in prior research in the field [34]. Removing flagged observations eliminated 77,684 observations, resulting in a dataset containing 2,429,283 observations. The latter were used in the analysis after imputing with random forest. The missing value profiles of the variables are shown below (Figures 1-5).

Figure 1. Missing value profile of the structural quality measures: surveys.

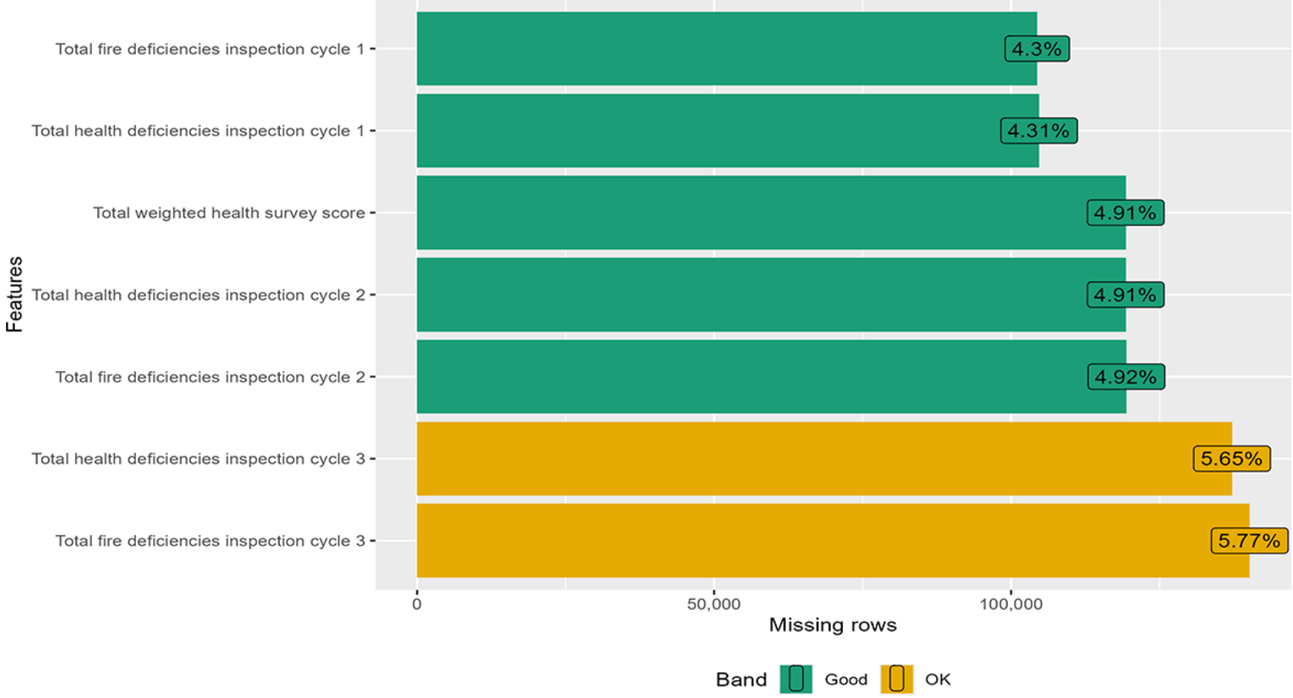


Figure 2. Missing value profile of the structural quality measures: provider information.

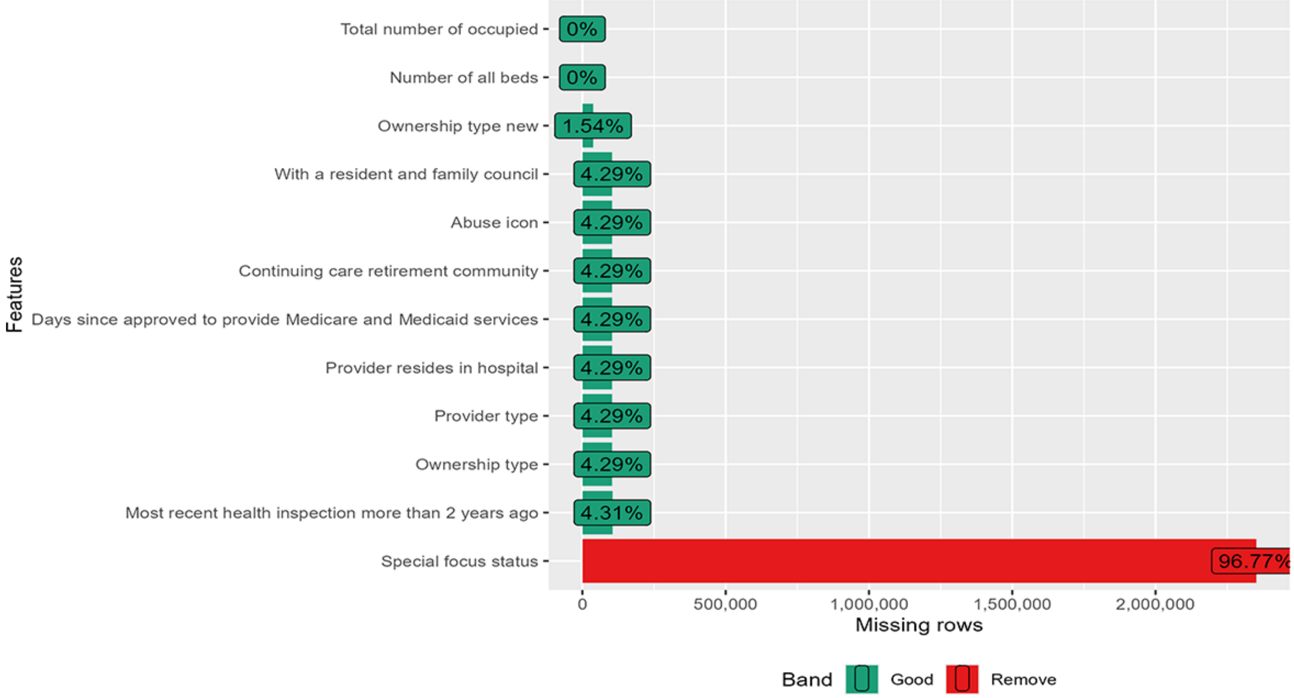


Figure 3. Missing value profile of the structural quality measures: penalties and staffing.

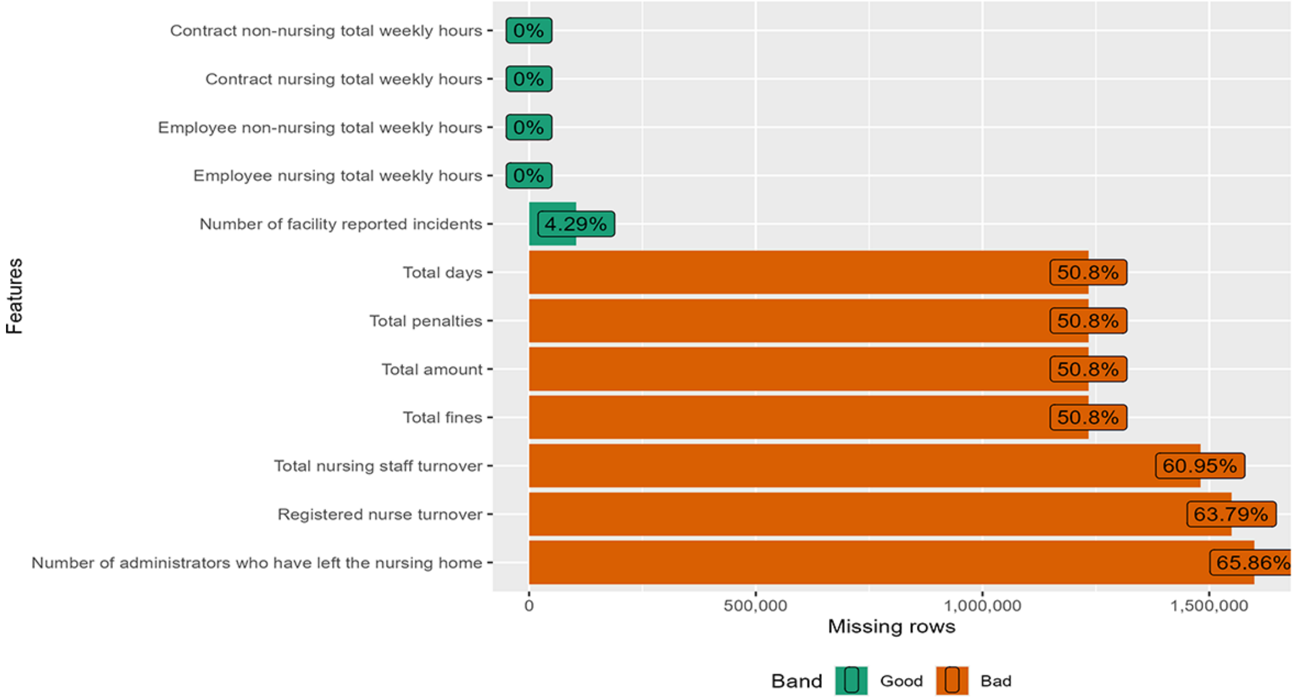


Figure 4. Missing value profile of the process quality measures.

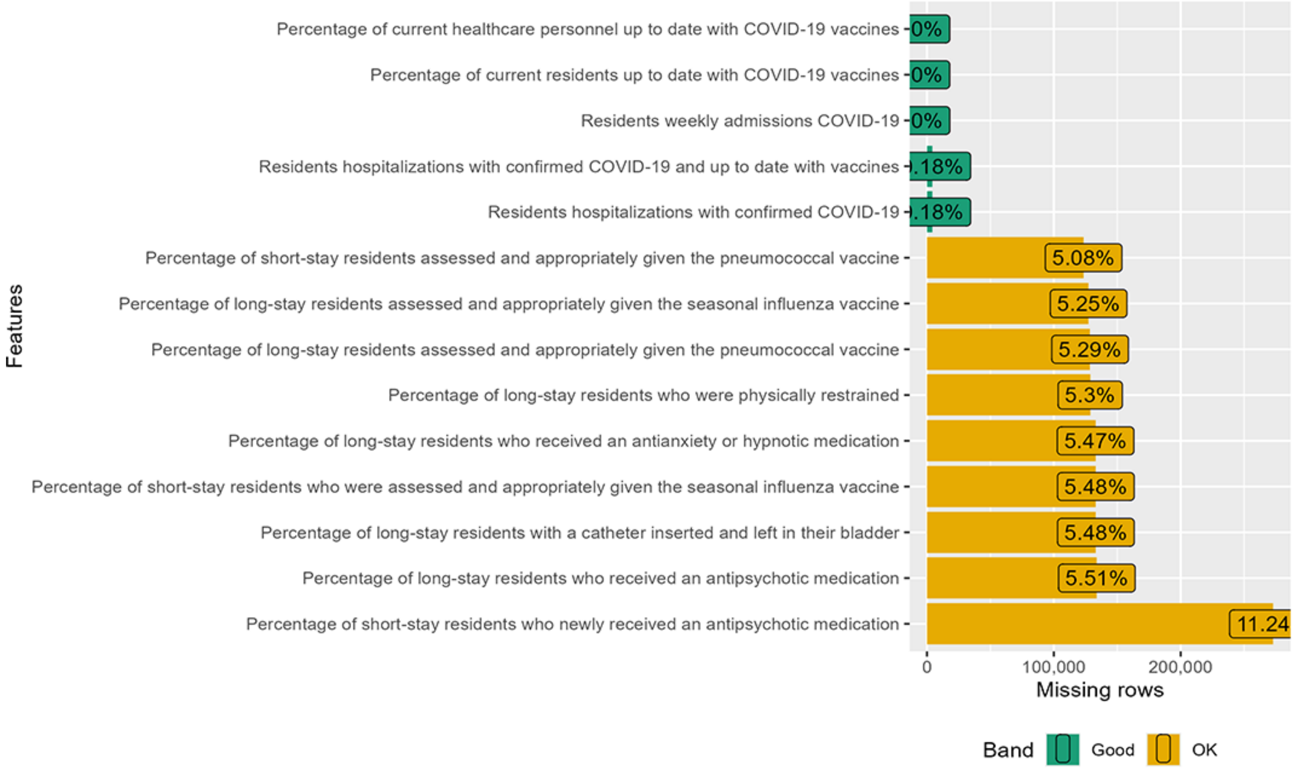
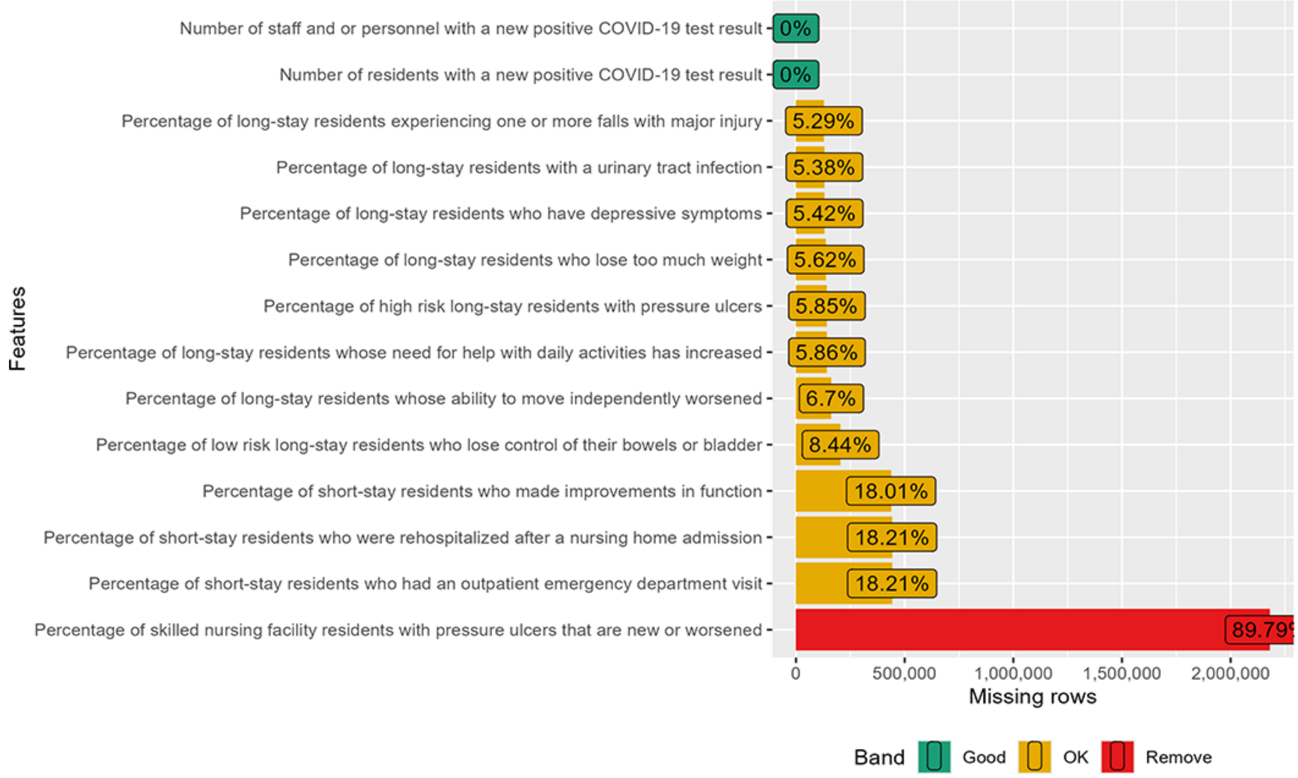


Figure 5. Missing value profile of the outcome quality measures.



Missing Value Imputation

When addressing the presence of missing values in our dataset, we made the deliberate choice to impute rather than exclude them, which aligned with the research goals and the nature of the missing data. In our study, we assumed that the missing data were missing at random (MAR). The MAR mechanism assumes that the probability of missingness depends on the observed data but not on the missing data [35]. As the nursing home data are aggregated at the institution level and include all institutions nationwide, missingness is likely related to differences in administrative or operational factors, such as staffing levels, documentation practices, reporting systems, or governance policies at the state, county, or local levels. We opted for random forest imputation, as it handles various data types seamlessly; it also excels in managing complex interactions and multicollinearity while efficiently scaling to accommodate large datasets. The implementation of random forest in the R package *missranger* [36] resulted in a reduction of up to 50% of imputation errors, and the average proportional deviation in out-of-bag imputation errors remained within the 10% to 15% range. All the variables (except the outcome variables) were imputed using the random forest algorithm. Although the best method to address MAR missingness is to perform multiple imputation [37], we were unable to adopt this approach due to time and resource constraints. The specifications used for random forest imputation are presented in Table 1.

Table . Specifications used for random forest imputation.

Parameter	Value
Predictive mean matching	5
Maximum iterations	20
Number of trees	20
Maximum depth	8

Data Modeling

Data Modeling

Using the imputed dataset, the variable called “percentage of occupied beds” was created using “total number of occupied beds” and “number of all beds.” A total of 10 observations from the calculation resulted in a not-a-number error and were excluded. All the numeric variables were standardized to have a mean of 0 and an SD of 1. All the unordered categorical variables or nominal variables were specified using sum contrasts. For ordered categorical variables or ordinal variables, we used orthogonal polynomial contrasts, and we standardized them onto a uniform scale. Consequently, each contrast column had a mean of 0 and an SD of 1. This approach allowed us to fit the regression models within a standardized framework and easily compare different model coefficients.

Furthermore, all unordered categorical variables containing more than 2 levels were binary encoded before being specified using sum contrasts. Binary encoding is a machine learning technique that combines hash encoding and one-hot encoding. First, all levels of an unordered categorical variable are expressed using a unique numeric value. Then, the numbers are transformed into a unique combination of zeros and ones, making it efficient to store data that exhibit high cardinality.

The outcome variables were characteristic of count data and exhibited properties such as nonnegativity, integer values, positive skewness, and a greater prevalence of lower count values. Equidispersion assessment revealed significant overdispersion, which rendered the use of a traditional Poisson regression model unsuitable. Therefore, we used a negative binomial regression model. Unlike the Poisson model, the negative binomial model accounts for overdispersion with quadratic parameterization. The presence of excess zeros led to the implementation of a zero-inflated negative binomial (ZINB) model, and the panel structure of the dataset resulted in the specification of a mixed effects model. However, specifying a ZINB model requires the presence of zero inflation at all levels of the grouping variables (ie, states, counties, and facilities). Some of the levels did not have zero inflation; they had zero deflation. On the basis of recommendations in the literature [38], the final model we specified was a truncated ZINB mixed effects model.

We estimated multiple truncated ZINB models, specifying different outcome variables. To model the outcomes “residents weekly confirmed COVID-19” and “staff weekly confirmed COVID-19,” the zero inflation (ZI) component of the model included factors influencing initial susceptibility to COVID-19 [39]. These included facility-level characteristics, such as bed size, as well as social determinants of health attributes, including urbanicity, political affiliation of the geographic location, median household income, and local COVID-19 vaccination coverage. To model the outcome “total weekly COVID-19 deaths,” the ZI component of the model included factors influencing the possibility of death after COVID-19 had been contracted, such as lagged weekly resident confirmed COVID-19 cases per 1000 residents and lagged county confirmed COVID-19 cases per 1000 people. We assumed that newly infected residents (ie, infected at 1 time point before) had a higher likelihood of death compared to those who were not newly infected (ie, infected more than one time point before). Our assumption was informed by prior research, indicating that individuals who test positive for COVID-19 typically do not begin to develop detectable antibodies until approximately 1 week after the onset of symptoms [40]. Both lagged variables were calculated for a lag of 1. An overview of the outcome variables and their associated ZI model predictors is presented in (Textbox 4).

Textbox 4. Model outcomes and ZI model predictors. Letter “N” denotes a numeric variable, whereas “C” denotes a categorical variable. The number preceding “C” represents the number of levels in the categorical variable.

Residents weekly confirmed COVID-19

- Bed size (3C)
- Urban binary (2C)
- Democrat or Republican (2C)
- Median household income dollars inflation adjusted to data file year ACS (American Community Survey) 2016-2020 (N)
- Partially or fully vaccinated percent (N)

Staff weekly confirmed COVID-19

- Bed size (3C)
- Urban binary (2C)
- Democrat or Republican (2C)
- Median household income dollars inflation adjusted to data file year ACS 2016-2020 (N)
- Partially or fully vaccinated percent (N)

Total weekly COVID-19 deaths

- Lagged weekly resident confirmed COVID-19 cases per 1000 residents (N)
- Lagged county confirmed cases USA Facts new (N)

The model included random slopes and random intercepts for 3 levels of nesting, that is, nursing home facilities nested within counties nested within states. An autoregressive process of order-1 covariance structure was selected to model the correlations among time points for the same individual. The truncated negative binomial distribution with quadratic parameterization implemented in the R package *glmmTMB* [41] under the family “truncated_nbinom2” was used.

Mathematically, the model can be expressed as follows:

$$\mu=E(count \mid u,NSZ)=\exp\ (\beta0+\beta ixi+u)\ \forall\ count>0$$

$$u\sim N(0,\sigma u2)$$

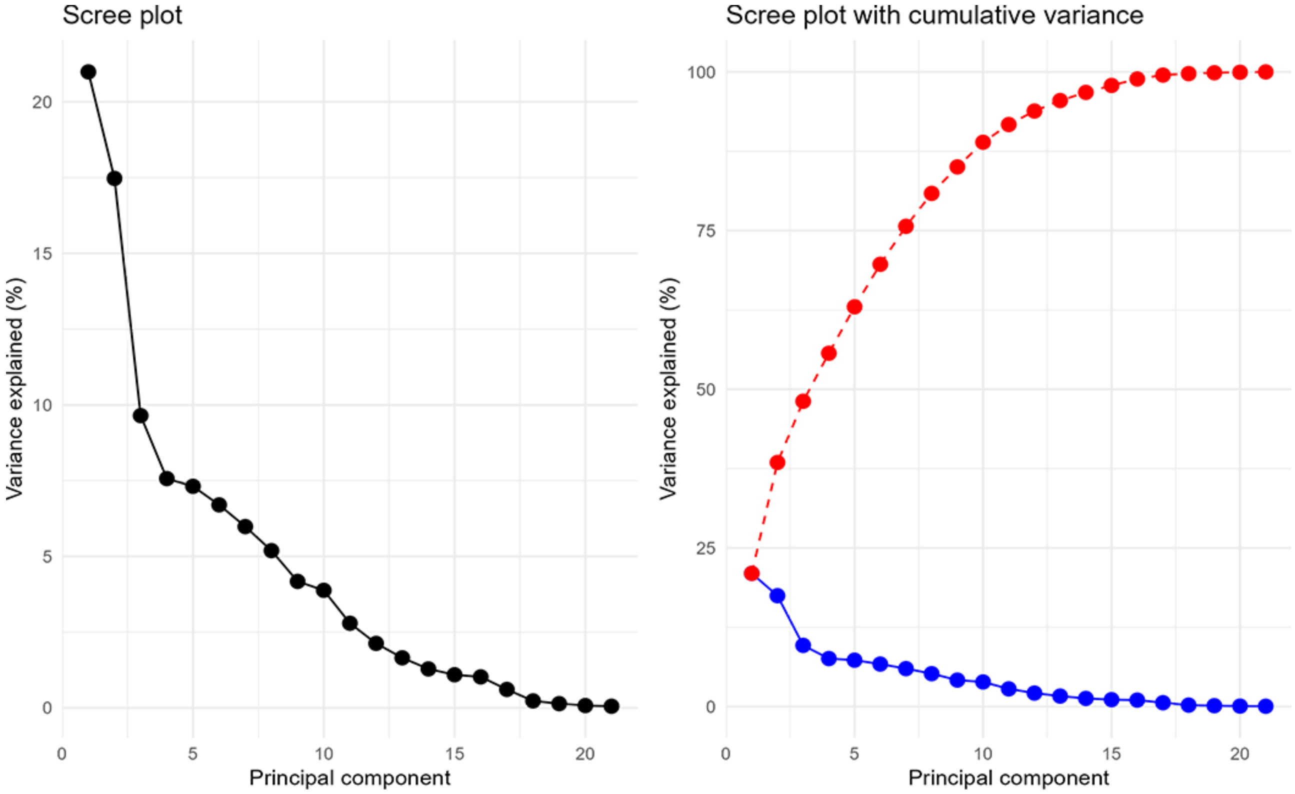
$$\sigma 2=Var(count|u,\ NSZ)=\ \mu(1+\mu\theta)\forall\ count>0$$

$$logitp=logp1-p=\beta0(zi)+\beta i(zi)xi$$

where subscript I denotes the time point, u denotes the random effects terms in the model, NSZ denotes the event “nonstructural zeros," *P*=1−Pr(NSZ) is the ZI probability, and θ is the dispersion parameter specific to the family "nbinom2."

For each outcome variable, we created 5 base models. Each model contained a different subset of nursing home QMs, namely, outcome QMs, process QMs, and structural QMs (surveys, provider information, and penalties and staffing). Using a subset of significant predictors (30/58) from the 5 base models, we performed principal component analysis (PCA) for dimensionality reduction on the mean-aggregated dataset, where data were averaged across repeated measures for each entity. Using the loadings (eigenvectors) from this PCA, we calculated principal component scores for the complete dataset, including all the repeated measures. Using the PCA scores as predictors, we created 2 predictive models for the states and the components of the 5-Star Quality Rating System of nursing homes. The first 8 principal components explained 80.86% of the variation present in 30 significant predictors (Figure 6).

Figure 6. Variation of the principal components.



Model Diagnostics

We performed model fit assessment through a combination of graphical diagnostics, goodness-of-fit statistics, and residual analysis. First, observed versus predicted count distributions were compared using histograms to visually assess the model's ability to accurately predict both zero and nonzero values. Second, goodness-of-fit statistics were computed, including the log-likelihood, Akaike Information Criterion, Bayesian Information Criterion, and deviance to evaluate and compare across statistical models. We selected several goodness-of-fit measures to achieve a more robust evaluation of model performance. Third, the Vuong test was applied to compare the ZINB model against standard negative binomial and Poisson models to determine whether the inclusion of a ZI component was statistically justified. Finally, residual diagnostics were performed using the R package *DHARMa* [42], and goodness-of-fit was evaluated through nonparametric, simulation-driven methods. *DHARMa* provided an assessment of outliers, over- or underdispersion, ZI, and the independence of observations by simulating new response data, computing empirical cumulative density functions, and defining residuals based on observed versus simulated values, ensuring a robust model adequacy assessment.

Ethical Considerations

This study is exempt from the institutional review board, as it involved only the use of publicly available data and did not involve human participants, in accordance with 45 CFR 46.102(e). No private, identifiable information was collected or analyzed.

Results

For the models involving outcome QMs as predictors, the number of new COVID-19 cases among residents (β=0.46, SE 0.002; β=0.02, SE 0. 001; and β=0.02, SE 0.003) and staff (β=0.01, SE 0.001; β=0.51, SE 0.002; and β=0.05, SE 0.004) were highly significant, with all the outcomes at the significance levels of *P*<.001 and *P*<.001 respectively. The percentage of long-stay residents with UTIs was significant, with staff infections as the outcome (β=−0.004, SE 0.002) at the significance level of *P*<.10 (*P*=0.06), resident infections as the outcome (β=−0.01, SE 0.003) at the significance level of *P*<.001, and total deaths as the outcome (β=−0.03, SE 0.009) at the significance level of *P*<.01 (*P*=.001).

For the models involving process QMs as predictors, the indicator concerning residents' weekly COVID-19 admissions (β=0.05, SE 0.003; β=0.04, SE 0.001; and β=0.048, SE 0.002) was significant, with all the outcomes at the significance level of *P*<.001.

Residents' hospitalizations with confirmed COVID-19 (β=0.03, SE 0.002 and β=0.02, SE 0.001), the percentage of current residents up to date with COVID-19 vaccines (β=0.06, SE 0.007 and β=0.02, SE 0.005), and the percentage of long-stay residents assessed and appropriately given the seasonal influenza vaccine (β=0.02, SE 0.005 and β=0.04, SE 0.01) were significant, with at least two of the outcomes at the significance levels of *P*<.001, *P*<.001, and *P*<.001, respectively.

Residents hospitalized with confirmed COVID-19 and up to date with vaccines (β=0.03, SE 0.005 and β=0.05, SE 0.02) and the percentage of long-stay residents with catheters inserted and left in their bladders (β=−0.05, SE 0.005; β=−0.01, SE 0.003; and β=−0.02, SE 0.009) were significant, with at least two of the outcomes at the significance level of *P*<.001 and *P*<.05, respectively.

For the models involving structural QMs (surveys) as predictors, total health deficiencies inspection cycle 2 (β=0.05, SE 0.005 and β=0.01, SE 0.004) was significant, with at least two outcomes at the *P*<.001 significance level.

Total health deficiencies inspection cycle 1 (β=0.04, SE 0.01 and β=0.01, SE 0.004), total health deficiencies inspection cycle 3 (β=0.04, SE 0.01 and β=0.01, SE 0.004), and total fire safety deficiencies inspection cycle 3 (β=0.02, SE 0.01 and β=0.01, SE 0.004) were significant, with at least one of the outcomes at the significance level of *P*<.001, *P*<.05, and *P*<.05, respectively.

For the models involving structural QM (provider information) as predictors, percentage of occupied beds (β=0.04, SE 0.005; β=−0.06, SE 0.004; and β=−0.25, SE 0.01), provider resides in hospital (β=0.19, SE 0.02; β=0.07, SE 0.01; and β=0.13, SE 0.03), and days since approval to provide Medicare and Medicaid services (β=0.08, SE 0.005; β=0.02, SE 0.005; and β=0.04, SE 0.01) were significant, with all outcomes at the *P*<.001, *P*<.001, and *P*<.001 significance levels, respectively.

Continuing care retirement community (β=0.08, SE 0.009 and β=0.06, SE 0.02) and at least one of the 2 binary components of provider type (component 1: β=0.26, SE 0.03; β=0.08, SE 0.03; and β=0.25, SE 0.06), with a resident and family council (component 1: β=0.02, SE 0.005; β=0.05, SE 0.01 and component 2: β=0.08, SE 0.01; β=0.03, SE 0.01; and β=0.16, SE 0.03), and new ownership type (component 2: β=0.05, SE 0.01 and β=−0.07, SE 0.01) were significant, with at least two of the outcomes at the *P*<.001, *P*<.001, *P*<.001, *P*<.001, and *P*<.001 significance level.

For the models involving structural QMs (penalties and staffing) as predictors, all the staffing variables, namely employee nurse total weekly hours (β=0.12, SE 0.006; β=0.09, SE 0.005; and β=0.06, SE 0.01), employee nonnursing total weekly hours (β=−1.06, SE 0.006; β=−0.09, SE 0.005; and β=−0.05, SE 0.01), contract nurse total weekly hours (β=0.1, SE 0.004; β=0.08, SE 0.003; and β=0.10, SE 0.008), and contract nonnursing total weekly hours (β=−0.1, SE 0.01; β=−0.1, SE 0.004; and β=−0.04, SE 0.01) were significant, with all the outcomes at the *P*<.001 significance level, respectively.

Component 3 of provider state (component 3: β=−0.23, SE 0.05; β=−0.23, SE 0.05; and β=−0.20, SE 0.06) was significant at the *P*<.001 significance level. None of the components of the provider county achieved a high level of significance.

The linear (β=0.08, SE 0.008 and β=0.05, SE 0.006), quadratic (β=0.02, SE 0.005 and β=0.02, SE 0.004), and cubic (β=0.03, SE 0.005 and β=0.02, SE 0.003) components of the health inspection rating were significant, with resident infections and staff infections as outcomes at the *P*<.001, *P*<.001, and *P*<.001 significance levels, respectively. However, only the quartic (β=−0.0156, SE 0.0079) component of the health inspection rating was significant, with total deaths as the outcome at the *P*<.05 (*P*=.048) significance level.

The QM rating had significant linear (β=−0.02, SE 0.01) and cubic components (β=−0.01, SE 0.005) in the model, with resident infections as the outcome at the *P*<.05 (*P*=.002) and *P*<.05 (*P*=.045) significance levels, respectively. In addition to the linear (β=0.0164, SE 0.0055) and cubic (β=−0.009, SE 0.003) components, the quartic (β=−0.006, SE 0.003) component was significant at the *P*<.05 (*P*=.003), *P*<.05 (*P*=.01), and *P*<.05 (*P*=.04) significance levels, respectively, in the model with staff infections as the outcome. The quadratic (β=−0.02, SE 0.01) and cubic (β=0.02, SE 0.01) components were significant at the *P*<.05 (*P*=.04) and *P*<.05 (*P*=.045) significance levels, respectively, in the model with total deaths as the outcome.

Staffing rating had a significant linear (β=0.02, SE 0.005) component in the models with staff infections and total deaths as the outcomes at the *P*<.001 significance level. The model with resident infections as the outcome only showed a significant cubic (β=−0.008, SE 0.004) trend at the *P*<.1 (*P*=.06) significance level. In the model with total deaths as the outcome, both the linear (β=0.12, SE 0.01) and quadratic (β=−0.05, SE 0.01) components were significant at the *P*<.001 and *P*<.001 significance levels, respectively, and the quartic (β=−0.01, SE 0.007) component was significant at the *P*<.1 (*P*=.05) significance level.

The discussed results are summarized in Tables 2 and 3. Multimedia Appendix 1, containing the complete model results in the original (log) scale, and Multimedia Appendix 2, containing the visualization of their 95% CIs in the response (exponent) scale, are provided for further reference.

Table . Model results for base models. The table displays variables that were significant when regressed on more than 1 outcome. All estimates are presented on the original (log) scale and are conditioned on the variables included in the conditional model and the model random effects. Complete results are provided in Multimedia Appendix 1 and displayed graphically in Multimedia Appendix 2.

Variable	Model 1: resident infections, estimate (SE)	*P* value	Model 2: staff infections, estimate (SE)	*P* value	Model 3: total deaths, estimate (SE)	*P* value
Intercept	4.60 (0.05)	<.001	3.52 (0.04)	<.001	3.89 (0.06)	<.001
Number of residents with a new positive COVID-19 test result	0.46 (0.002)	<.001	0.02 (0.001)	<.001	0.02 (0.003)	<.001
Number of staff and/or personnel with a new positive COVID-19 test result	0.01 (0.001)	<.001	0.51 (0.002)	<.001	0.05 (0.004)	<.001
Percentage of long-stay residents whose ability to move independently worsened	0.01 (0.003)	.02	0.01 (0.003)	.007	−0.02 (0.01)	.22
Percentage of long-stay residents with a urinary tract infection	−0.02 (0.003)	<.001	−0.004 (0.002)	.06	−0.03 (0.01)	.001
Intercept	4.28 (0.08)	<.001	3.31 (0.06)	<.001	3.37 (0.06)	<.001
Residents weekly admissions COVID-19	0.05 (0.003)	<.001	0.04 (0.001)	<.001	0.05 (0.002)	<.001
Residents hospitalizations with confirmed COVID-19	0.03 (0.002)	<.001	0.02 (0.001)	<.001	0.01 (0.004)	.26
Residents hospitalizations with confirmed COVID-19 and up to date with vaccines	0.0003 (0.002)	.92	0.03 (0.005)	<.001	0.05 (0.02)	.001
Percentage of current residents up to date with COVID-19 vaccines	−0.06 (0.01)	<.001	−0.02 (0.005)	<.001	−0.10 (0.03)	.002
Percentage of current health care personnel up to date with COVID-19 vaccines	0.03 (0.01)	<.001	0.01 (0.004)	.04	−0.001 (0.03)	.99
Percentage of long-stay residents assessed and appropriately given the seasonal influenza vaccine	0.02 (0.01)	<.001	0.0001 (0.004)	.97	0.04 (0.01)	<.001
Percentage of long-stay residents who received an antipsychotic medication	0.04 (0.01)	<.001	0.01 (0.005)	.13	0.2 (0.1)	.06
Percentage of long-stay residents who were physically restrained	−0.02 (0.01)	<.001	−0.002 (0.004)	.70	.02 (0.01)	.09

Variable	Model 1: resident infections, estimate (SE)	*P* value	Model 2: staff infections, estimate (SE)	*P* value	Model 3: total deaths, estimate (SE)	*P* value
Percentage of long-stay residents with a catheter inserted and left in their bladder	−0.05 (0.005)	<.001	−0.01 (0.003)	.04	−0.02 (0.01)	.009
Percentage of short-stay residents who were assessed and appropriately given the seasonal influenza vaccine	−0.04 (0.01)	<.001	0.001 (0.01)	.85	−0.0004 (0.01)	.03
Intercept	4.70 (0.08)	<.001	3.53 (0.06)	<.001	3.86 (0.06)	<.001
Total health deficiencies inspection cycle	0.04 (0.01)	<.001	0.01 (0.005)	.009	0.02 (0.01)	.09
Total health deficiencies inspection cycle 2	0.05 (0.01)	<.001	0.01 (0.004)	<.001	0.01 (0.01)	.30
Total health deficiencies inspection cycle 3	0.04 (0.01)	<.001	0.01 (0.004)	.007	0.02 (0.01)	.10
Total fire deficiencies inspection cycle 3	0.02 (0.005)	<.001	0.01 (0.004)	.006	0.02 (0.01)	.06
Intercept	3.98 (0.08)	<.001	3.35 (0.06)	<.001	3.13 (0.09)	<.001
Percent of occupied beds	0.04 (0.01)	<.001	−0.06 (0.004)	<.001	−0.25 (0.01)	<.001
Provider type b1	0.26 (0.03)	<.001	0.08 (0.03)	.001	0.25 (0.06)	<.001
Provider type b2	−0.02 (0.02)	.07	−0.001 (0.02)	.01	−0.03 (0.05)	.09
Provider resides in hospital	0.18 (0.02)	<.001	0.07 (0.01)	<.001	0.13 (0.03)	<.001
Days since approved to provide Medicare and Medicaid services	0.08 (0.01)	<.001	0.02 (0.005)	<.001	0.04 (0.01)	<.001
Continuing care retirement community	0.08 (0.01)	<.001	−0.01 (0.01)	.43	0.06 (0.02)	<.001
Special focus status	−0.03 (0.004)	<.001	−0.003 (0.004)	.07	−0.01 (0.01)	.29
With a resident and family council b1	0.01 (0.01)	.11	0.02 (0.01)	<.001	0.05 (0.01)	<.001
With a resident and family council b2	0.08 (0.01)	<.001	0.03 (0.01)	.001	0.16 (0.03)	<.001
Ownership type new b1	0.01 (0.01)	<.001	0.02 (0.01)	.06	0.04 (0.01)	.002
Ownership type new b2	0.05 (0.01)	<.001	−0.07 (0.01)	<.001	−0.03 (0.02)	.14
Intercept	4.70 (0.08)	<.001	3.52 (0.06)	<.001	3.85 (0.06)	<.001
Total days	0.01 (0.01)	.08	0.01 (0.004)	.002	0.004 (0.01)	.72
Employee nursing total weekly hours	0.12 (0.01)	<.001	0.09 (0.01)	<.001	0.06 (0.01)	<.001

Variable	Model 1: resident infections, estimate (SE)	*P* value	Model 2: staff infections, estimate (SE)	*P* value	Model 3: total deaths, estimate (SE)	*P* value
Employee nonnursing total weekly hours	−0.11 (0.01)	<.001	−0.09 (0.05)	<.001	−0.05 (0.01)	<.001
Contract nursing total weekly hours	0.10 (0.004)	<.001	0.08 (0.003)	<.001	0.10 (0.01)	<.001
Contract non-nursing total weekly hours	−0.10 (0.01)	<.001	−0.09 (0.004)	<.001	−0.04 (0.01)	<.001
Number of facility reported incidents	0.02 (0.01)	<.001	−0.001 (0.004)	.80	0.02 (0.01)	.06
Registered nurse turnover	0.01 (0.005)	.002	−0.01 (0.003)	.008	0.01 (0.01)	.09

Table . Model results for prediction models. The table displays variables that were significant when regressed on more than one outcome. All estimates are presented on the original (log) scale and are conditioned on the variables included in the conditional model and the model random effects. Complete results are provided in Multimedia Appendix 1 and displayed graphically in Multimedia Appendix 2.

Variable	Model 1: resident infections, estimate (SE)	*P* value	Model 2: staff infections, estimate (SE)	*P* value	Model 3: total deaths, estimate (SE)	*P* value
Intercept rating	4.36 (0.05)	<.001	2.97 (0.05)	<.001	3.45 (0.06)	<.001
PC[a] rating	0.26 (0.01)	<.001	0.17 (0.004)	<.001	0.04 (0.01)	<.001
PC2 rating	0.59 (0.005)	<.001	0.53 (0.003)	<.001	0.02 (0.01)	<.001
PC3 rating	−0.06 (0.01)	<.001	−0.09 (0.004)	<.001	−0.01 (0.009)	.06
PC4 rating	0.12 (0.005)	<.001	0.16 (0.003)	<.001	−0.02 (0.01)	.003
PC5 rating	−0.005 (0.01)	.42	−0.06 (0.005)	<.001	0.03 (0.01)	.005
PC6 - rating	0.17 (0.01)	<.001	0.3 (0.04)	<.001	0.21 (0.01)	<.001
PC7 - rating	0.06 (0.01)	<.001	0.08 (0.004)	<.001	0.25 (0.01)	<.001
PC8 - rating	−0.09 (0.005)	<.001	−0.2 (0.004)	<.001	−0.08 (0.01)	<.001
Health inspection rating.l	0.08 (0.01)	<.001	0.05 (0.01)	<.001	0.02 (0.01)	.16
Health inspection rating.q	0.02 (0.01)	<.001	0.02 (0.004)	<.001	0.004 (0.01)	.67
Health inspection rating.c	0.03 (0.005)	<.001	0.02 (0.003)	<.001	0.01 (0.01)	.05
QM rating.l	−0.02 (0.01)	.002	0.02 (0.01)	.003	0.01 (0.01)	.54
QM rating.c	−0.01 (0.005)	.04	−0.01 (0.003)	.01	0.02 (0.01)	.05
Staffing rating.l	−0.01 (0.01)	.13	0.02 (0.005)	<.001	−0.12 (0.01)	<.001
Intercept - provider state	4.37 (0.06)	<.001	2.98 (0.06)	<.001	3.45 (0.07)	<.001
PC - provider state	−0.24 (0.004)	<.001	−0.15 (0.003)	<.001	−0.05 (0.01)	<.001
PC2 - provider state	0.59 (0.004)	<.001	0.53 (0.003)	<.001	0.03 (0.01)	<.001
PC3 - provider state	−0.08 (0.01)	<.001	−0.1 (0.004)	<.001	−0.01 (0.008)	.05
PC4 - provider state	−0.12 (0.005)	<.001	−0.16 (0.003)	<.001	0.02 (0.01)	<.001
PC5 - provider state	0.005 (0.01)	.74	−0.05 (0.005)	<.001	0.03 (0.01)	<.001
PC6 - provider state	0.16 (0.01)	<.001	0.29 (0.004)	<.001	0.19 (0.01)	<.001
PC7 - provider state	−0.08 (0.01)	<.001	−0.09 (0.004)	<.001	−0.26 (0.01)	<.001
PC8 - provider state	−0.09 (0.005)	<.001	−0.2 (0.004)	<.001	−0.07 (0.01)	<.001
Provider state b3	−0.23 (0.05)	<.001	−0.23 (0.05)	<.001	−0.2 (0.06)	<.001
Provider state b4	−0.01 (0.05)	.01	−0.03 (0.05)	.55	0.11 (0.06)	.07
Intercept - provider county	4.31 (0.07)	<.001	2.96 (0.07)	<.001	3.47 (0.08)	<.001
PC - provider county	−0.24 (0.004)	<.001	−0.15 (0.003)	<.001	−0.05 (0.01)	<.001
PC2 - provider county	0.59 (0.004)	<.001	0.53 (0.003)	<.001	0.03 (0.01)	<.001
PC3 - provider county	−0.08 (0.01)	<.001	−0.1 (0.004)	<.001	−0.01 (0.008)	.09

Variable	Model 1: resident infections, estimate (SE)	*P* value	Model 2: staff infections, estimate (SE)	*P* value	Model 3: total deaths, estimate (SE)	*P* value
PC4 - provider county	−0.12 (0.005)	<.001	−0.16 (0.003)	<.001	0.02 (0.01)	<.001
PC5 - provider county	0.004 (0.01)	.52	−0.05 (0.005)	<.001	0.03 (0.01)	<.001
PC6 - provider county	0.16 (0.01)	<.001	0.29 (0.004)	<.001	0.19 (0.01)	<.001
PC7 - provider county	−0.08 (0.01)	<.001	−0.09 (0.004)	<.001	−0.26 (0.01)	<.001
PC8 - provider county	−0.09 (0.005)	<.001	−0.2 (0.004)	<.001	−0.07 (0.01)	<.001

[a] PC: principal component.

Discussion

Principal Findings

In this study, we used a panel dataset to evaluate the effectiveness of several nursing home QMs in explaining nursing home QoC in terms of COVID-19 outcomes. Staffing measures (*P*<.001 for all variables in all models), health deficiency scores (*P*<.001 for all variables in at least 1 model), COVID-19 hospitalizations (*P*<.001 for all variables in at least 2 models), and vaccinations (*P*<.001 for all variables in at least 2 models) exhibited meaningful relationships with the COVID-19 outcomes. The 5-Star Quality Rating System for nursing homes, Medicaid dependency, and ownership showed no clear relationships with the COVID-19 outcomes.

We found a significant association with the percentage of long-stay residents with UTIs. In a retrospective study, the authors evaluated UTI diagnoses and antibiotic prescriptions in 622 COVID-19 hospital ward patients, and they found that 61% of cases had probably been overdiagnosed [43]. Several researchers studying long-term care facilities have also found evidence of overdiagnosis in the form of inappropriate initiation of antimicrobial treatment in 37% to 61% of patients [43-49].

We also identified significant associations with the percentages of long-stay and short-stay residents assessed and appropriately given the seasonal influenza vaccine. A study compared COVID-19 and seasonal influenza patient groups and suggested that preexisting chronic respiratory conditions more strongly impacted the severity of seasonal influenza than that of COVID-19 [50]. It has also been reported that chronic lower respiratory tract diseases claim the third highest number of lives among adults aged >65 years [51-55]. Therefore, older individuals with comorbidities for both respiratory infections may exhibit stronger symptoms of seasonal influenza than COVID-19, demanding management through the administration of seasonal influenza vaccines.

A significant association was uncovered for the percentage of long-stay residents with catheters inserted and left in their bladders. In a recent systematic review of 67 studies, scholars found that catheter prevalence among nursing home residents varied between 2.2% and 36.4%, with a typical rate of 7.3% [56,57]. Catheter-associated UTIs comprise 32% of all health care–associated infections [57,58]. The fact that catheter use is frequently reported in long-term care residents and is a leading cause of UTI helps explain our results, which show significant associations with COVID-19 outcomes for both the percentage of long-stay residents with UTIs and the percentage of long-stay residents with catheters inserted and left in their bladders.

Our findings indicate a significant association between health deficiency variables and resident and staff COVID-19 cases. During the period 2013 to 2017, more than four-fifths of nursing homes in the United States engaged in deficient infection prevention and control practices, including those that substantially reduced the spread of COVID-19, such as proper hand washing and isolation procedures [21,59]. It is possible that these poor practices continued during the pandemic, making nursing home residents and staff more susceptible to COVID-19 infections.

Several scholars have reported that lower resident density, measured by the proxy variable of the percentage of occupied beds, is associated with a lower prevalence of COVID-19 [59-65]. Our study showed a similar association with resident infections but an inverse association with staff infections and total deaths. We found that hospital-based nursing homes reported fewer COVID-19 cases and deaths than nonhospital-based nursing homes, which is consistent with the results of a study by Tarteret et al [66]. Gorges and Konetzka [67] demonstrated facility-level differences in QoC based on the proportion of residents with Medicaid coverage at the facility. Compared to nursing homes accepting Medicaid only, nursing homes accepting Medicare showed clearly better COVID-19 outcomes in only 2 of 3 models. In 1 model, the difference was less clear due to overlapping 95% CIs.

Evidence from various studies suggests that for-profit nursing homes are more prone to COVID-19 outbreaks compared to nonprofit and government nursing homes [59,61,65,68]. However, some scholars have found no statistically significant link in this regard [34,60,66]. We did not notice a clear difference between for-profit, nonprofit, and government nursing homes. Furthermore, no clear difference in outcomes was observed between nursing homes with or without continuing care retirement communities, resident councils, family councils, or resident and family councils. While not a key characteristic, nursing homes newly approved to provide Medicare and Medicaid services had better COVID-19 outcomes compared to long-standing nursing homes. This could be because newer nursing homes underwent inspections more recently than previously approved homes as part of their approval process.

Inadequate nursing staff have been linked to a higher likelihood of COVID-19 outbreaks in some studies [59,62,64,65,68-70], but others indicate no significant connection [61,65,69,71]. Our findings show that increased staffing hours for nursing staff resulted in worse COVID-19 outcomes, whereas increased staffing hours for non-nursing staff resulted in better COVID-19 outcomes. We did not explore these associations further based on the types of nursing and nonnursing staff.

Most researchers have found no meaningful connection between nursing home overall quality rating and COVID-19 outcomes [34,60-62,69,72-79]. Similarly, most scholars have found no clear relationships between COVID-19 outcomes and 2 components of the 5-Star Quality Rating System for Nursing Homes, namely staffing rating and health inspection rating [61,64,70,73,78]. Our analysis did not reveal a clear relationship between the components of the 5-Star Quality Rating System for nursing homes and COVID-19 outcomes.

Limitations

This study has some limitations. It only considered temporal autocorrelation, without addressing spatial autocorrelation, thus ignoring potential interdependencies across spatial units. Temporal autocorrelation assumes that changes over time are isolated within units, which is rarely the case in interconnected systems, such as health care networks. We also did not perform multiple dataset imputations, which could have resulted in less robust statistical analyses, as a single imputation set may not adequately reflect the variability inherent in missing data. By generating a range of plausible datasets, MIs allow for more accurate and reliable parameter estimates by accounting for the uncertainty around the imputed values. Future scholars should consider modeling both spatial and temporal autocorrelations, as well as performing MIs, to increase the validity and reliability of their findings.

Conclusions

The evolution of quality in the nursing home industry is truly remarkable. However, sustained progress is essential to meet the needs of the growing older adult population. Although widely used in the nursing home industry, the 5-Star Quality Rating System for nursing homes is an unreliable performance measure. As demonstrated through the results of the study, it is hard to delineate nursing homes' performance based on the star ratings, after accounting for the underlying QMs. Concerted efforts from lawmakers, policy makers, and lobbyists will be required to refine and enhance the measure, thereby ensuring its reliability and effectiveness.

Funding

The authors did not receive support from any funding agency in the public, commercial, or not-for-profit sectors for the submitted work.

Data Availability

The datasets used in this study are available from the author's figshare repositories and are listed in the references.

Authors' Contributions

Conceptualization: HA
Data curation: SB
Formal analysis: SB
Writing – original draft: SB (lead), HA (supporting)
Writing – review & editing: SB (lead), HA (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete set of results tables for all models estimated in the study. Tables are organized by outcome variable and by model specification, distinguishing between base models using the original variables and final models incorporating principal component–derived variables, respectively.
[DOCX File, 102 KB - humanfactors_v13i1e72770_app1.docx]

Multimedia Appendix 2
Complete set of results displayed graphically. Bars denote the point estimates, while whiskers indicate the corresponding 95% CIs. Each figure is divided into 2 panels, representing the conditional model and the zero-inflation model, respectively.
[DOCX File, 1452 KB - humanfactors_v13i1e72770_app2.docx]

References

1. Nursing home care. National Center for Health Statistics. 2020. URL: https://www.cdc.gov/nchs/fastats/nursing-home-care.htm [accessed 2025-09-11]
2. Davidson PM, Szanton SL. Nursing homes and COVID-19: we can and should do better. J Clin Nurs 2020 Aug;29(15-16):2758-2759. [doi: 10.1111/jocn.15297] [Medline: 32281165]
3. McGarry BE, Gandhi AD, Barnett ML. COVID-19 surveillance testing and resident outcomes in nursing homes. N Engl J Med 2023 Mar 23;388(12):1101-1110. [doi: 10.1056/NEJMoa2210063] [Medline: 36947467]
4. Grabowski DC, Chen A, Saliba D. Paying for nursing home quality: an elusive but important goal. J American Geriatrics Society 2023 Feb;71(2):342-348. [doi: 10.1111/jgs.18260]
5. Luo H, Lum TYS, Wong GHY, Kwan JSK, Tang JYM, Chi I. Predicting adverse health outcomes in nursing homes: a 9-year longitudinal study and development of the FRAIL-Minimum Data Set (MDS) quick screening tool. J Am Med Dir Assoc 2015 Dec;16(12):1042-1047. [doi: 10.1016/j.jamda.2015.09.006] [Medline: 26602760]
6. Moore KL, Boscardin WJ, Steinman MA, Schwartz JB. Patterns of chronic co-morbid medical conditions in older residents of U.S. nursing homes: differences between the sexes and across the agespan. J Nutr Health Aging 2014 Apr;18(4):429-436. [doi: 10.1007/s12603-014-0001-y] [Medline: 24676326]
7. Reilev M, Lundby C, Jensen J, Larsen SP, Hoffmann H, Pottegård A. Morbidity and mortality among older people admitted to nursing home. Age Ageing 2019 Dec 1;49(1):67-73. [doi: 10.1093/ageing/afz136] [Medline: 31763682]
8. Vespa J, Armstrong DM, Medina L. Demographic turning points for the United States: population projections for 2020 to 2060. : United States Census Bureau; 2018 URL: https://www.census.gov/content/dam/Census/library/publications/2020/demo/p25-1144.pdf [accessed 2026-01-06]
9. Consumer Financial Protection Bureau Office for Older Americans. Issue spotlight: nursing home debt collection. : Consumer Financial Protection Bureau; 2022 URL: https://www.consumerfinance.gov/data-research/research-reports/issue-spotlight-nursing-home-debt-collection/full-report/ [accessed 2026-01-06]
10. Medina L, Sabo S, Vespa J. Living longer: historical and projected life expectancy in the United States, 1960 to 2060. : United States Census Bureau; 2020 URL: https://www.census.gov/content/dam/Census/library/publications/2020/demo/p25-1145.pdf [accessed 2026-01-06]
11. Mitzner TL, Beer JM, McBride SE, Rogers WA, Fisk AD. Older adults' needs for home health care and the potential for human factors interventions. Proc Hum Factors Ergon Soc Annu Meet 2009 Oct;53(1):718-722. [doi: 10.1177/154193120905301118] [Medline: 25530696]
12. Ruhm CJ. Excess deaths in the United States during the first year of COVID-19. Prev Med 2022 Sep;162:107174. [doi: 10.1016/j.ypmed.2022.107174] [Medline: 35878708]
13. Adams ML, Katz DL, Grandpre J. Population-based estimates of chronic conditions affecting risk for complications from coronavirus disease, United States. Emerg Infect Dis 2020 Aug;26(8):1831-1833. [doi: 10.3201/eid2608.200679] [Medline: 32324118]
14. Williams CS, Zheng Q, White AJ, et al. The association of nursing home quality ratings and spread of COVID-19. J Am Geriatr Soc 2021 Aug;69(8):2070-2078. [doi: 10.1111/jgs.17309] [Medline: 34058015]
15. COVID-19: who's at higher risk of serious symptoms. Mayo Clinic. 2024. URL: https://www.mayoclinic.org/diseases-conditions/coronavirus/in-depth/coronavirus-who-is-at-risk/art-20483301 [accessed 2025-09-11]
16. De Witte J, Noten S, Vinckers F, et al. "You needed to accept the situation": resilience of nursing home residents in times of COVID-19. Can J Aging 2024 Mar;43(1):75-83. [doi: 10.1017/S0714980823000399] [Medline: 37665004]
17. Rubano MD, Kieffer EF, Larson EL. Infection prevention and control in nursing homes during COVID-19: an environmental scan. Geriatr Nurs 2022;43:51-57. [doi: 10.1016/j.gerinurse.2021.10.023] [Medline: 34808421]
18. Dorritie R, Quigley DD, Agarwal M, Tark A, Dick A, Stone PW. Support of nursing homes in infection management varies by US State Departments of Health. J Hosp Infect 2020 Jun;105(2):258-264. [doi: 10.1016/j.jhin.2020.02.007] [Medline: 32068013]
19. Ferdous F. Social distancing vs social interaction for older adults at long-term care facilities in the midst of the COVID-19 pandemic: a rapid review and synthesis of action plans. Inquiry 2021;58:469580211044287. [doi: 10.1177/00469580211044287] [Medline: 34612063]
20. How to protect yourself and others. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/covid/prevention/index.html [accessed 2025-09-11]
21. Infection control deficiencies were widespread and persistent in nursing homes prior to COVID-19 pandemic. United States Government Accountability Office. 2020. URL: https://www.gao.gov/products/gao-20-576r [accessed 2025-09-11]
22. Nursing home quality initiative. : Centers for Medicare & Medicaid Services; 2003 URL: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Downloads/NHQIOverView20030731.pdf [accessed 2025-09-11]
23. CMS' five star quality rating system for nursing homes. Centers for Medicare & Medicaid Services. 2014. URL: https://www.cms.gov/newsroom/fact-sheets/cms-five-star-quality-rating-system-nursing-homes [accessed 2025-09-11]
24. Castle NG, Ferguson JC. What is nursing home quality and how is it measured? Gerontologist 2010 Aug;50(4):426-442. [doi: 10.1093/geront/gnq052] [Medline: 20631035]

25. Castle NG, Engberg J. Staff turnover and quality of care in nursing homes. Med Care 2005 Jun;43(6):616-626. [doi: 10.1097/01.mlr.0000163661.67170.b9] [Medline: 15908857]
26. Lyder CH. Pressure ulcer prevention and management. JAMA 2003 Jan 8;289(2):223-226. [doi: 10.1001/jama.289.2.223] [Medline: 12517234]
27. Clemens S, Wodchis W, McGilton K, McGrail K, McMahon M. The relationship between quality and staffing in long-term care: a systematic review of the literature 2008-2020. Int J Nurs Stud 2021 Oct;122:104036. [doi: 10.1016/j.ijnurstu.2021.104036] [Medline: 34419730]
28. Collier E, Harrington C. Staffing characteristics, turnover rates, and quality of resident care in nursing facilities. Res Gerontol Nurs 2008 Jul;1(3):157-170. [doi: 10.3928/19404921-20080701-02] [Medline: 20077960]
29. Spilsbury K, Hewitt C, Stirk L, Bowman C. The relationship between nurse staffing and quality of care in nursing homes: a systematic review. Int J Nurs Stud 2011 Jun;48(6):732-750. [doi: 10.1016/j.ijnurstu.2011.02.014] [Medline: 21397229]
30. Hillmer MP, Wodchis WP, Gill SS, Anderson GM, Rochon PA. Nursing home profit status and quality of care: is there any evidence of an association? Med Care Res Rev 2005 Apr;62(2):139-166. [doi: 10.1177/1077558704273769] [Medline: 15750174]
31. R Foundation for Statistical Computing. 2025. URL: https://www.R-project.org/ [accessed 2026-01-06]
32. McClelland GH. Increasing statistical power without increasing sample size. Am Psychol 2000;55(8):963-964. [doi: 10.1037//0003-066X.55.8.963]
33. Donabedian A. The quality of care: how can it be assessed? JAMA 1988;260(12):1743-1748. [doi: 10.1001/jama.1988.03410120089033]
34. Braun RT, Yun H, Casalino LP, et al. Comparative performance of private equity-owned US nursing homes during the COVID-19 pandemic. JAMA Netw Open 2020 Oct 1;3(10):e2026702. [doi: 10.1001/jamanetworkopen.2020.26702] [Medline: 33112402]
35. Newman DA. Missing data. Organ Res Methods 2014 Oct;17(4):372-411. [doi: 10.1177/1094428114548590]
36. Mayer M. MissRanger: fast imputation of missing values R package version 2.6.0. Comprehensive R Archive Network (CRAN). 2024. URL: https://CRAN.R-project.org/package=missRanger [accessed 2026-01-06]
37. Penny KI, Atkinson I. Approaches for dealing with missing data in health care studies. J Clin Nurs 2012 Oct;21(19-20):2722-2729. [doi: 10.1111/j.1365-2702.2011.03854.x] [Medline: 21895816]
38. Feng CX. A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. J Stat Distrib Appl 2021;8(1):8. [doi: 10.1186/s40488-021-00121-4] [Medline: 34760432]
39. Giri S, Chenn LM, Romero-Ortuno R. Nursing homes during the COVID-19 pandemic: a scoping review of challenges and responses. Eur Geriatr Med 2021 Dec;12(6):1127-1136. [doi: 10.1007/s41999-021-00531-2] [Medline: 34136990]
40. Interim guidelines for COVID-19 antibody testing. Centers for Disease Control and Prevention. URL: https://archive.cdc.gov/#/details?q=https://www.cdc.gov/coronavirus/2019-ncov/hcp/testing/antibody-tests-guidelines.html&start=0&rows=10&url=https://www.cdc.gov/coronavirus/2019-ncov/hcp/testing/antibody-tests-guidelines.html [accessed 2025-09-11]
41. Brooks M, Kristensen K, Benthem K, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J 2017;9(2):378 [FREE Full text] [doi: 10.32614/RJ-2017-066]
42. Hartig F. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models (version 046) [R package]. Comprehensive R Archive Network (CRAN). URL: https://CRAN.R-project.org/package=DHARMa [accessed 2025-09-11]
43. Van Laethem J, Wuyts SCM, Pierreux J, et al. Presumed urinary tract infection in patients admitted with COVID-19: are we treating too much? Antibiotics (Basel) 2021 Dec 6;10(12):1493. [doi: 10.3390/antibiotics10121493] [Medline: 34943705]
44. Lim CJ, McLellan SC, Cheng AC, et al. Surveillance of infection burden in residential aged care facilities. Med J Aust 2012 Mar 19;196(5):327-331 [FREE Full text] [doi: 10.5694/mja12.10085] [Medline: 22432671]
45. Loeb M, Simor AE, Landry L, et al. Antibiotic use in Ontario facilities that provide chronic care. J Gen Intern Med 2001 Jun;16(6):376-383. [doi: 10.1046/j.1525-1497.2001.016006376.x] [Medline: 11422634]
46. Montgomery P, Semenchuk M, Nicolle LE. Antimicrobial use in nursing homes in Manitoba. J Geriatr Drug Ther 1995;9(3):55-74. [doi: 10.1300/J089v09n03_05]
47. Peron EP, Hirsch AA, Jury LA, Jump RLP, Donskey CJ. Another setting for stewardship: high rate of unnecessary antimicrobial use in a veterans affairs long-term care facility. J Am Geriatr Soc 2013 Feb;61(2):289-290. [doi: 10.1111/jgs.12099] [Medline: 23405923]
48. Stuart RL, Wilson J, Bellaard-Smith E, et al. Antibiotic use and misuse in residential aged care facilities. Intern Med J 2012 Oct;42(10):1145-1149. [doi: 10.1111/j.1445-5994.2012.02796.x] [Medline: 22472087]
49. Zimmer JG, Bentley DW, Valenti WM, Watson NM. Systemic antibiotic use in nursing homes. A quality assessment. J Am Geriatr Soc 1986 Oct;34(10):703-710. [doi: 10.1111/j.1532-5415.1986.tb04301.x] [Medline: 3760435]
50. Brehm TT, van der Meirschen M, Hennigs A, et al. Comparison of clinical characteristics and disease outcome of COVID-19 and seasonal influenza. Sci Rep 2021 Mar 11;11(1):5803. [doi: 10.1038/s41598-021-85081-0] [Medline: 33707550]
51. Soriano JB, Kendrick PJ, Paulson KR. Prevalence and attributable health burden of chronic respiratory diseases, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Respir Med 2020 Jun;8(6):585-596. [doi: 10.1016/S2213-2600(20)30105-3] [Medline: 32526187]

52. Joshi PR. Pulmonary diseases in older patients: understanding and addressing the challenges. Geriatrics (Basel) 2024 Mar 7;9(2):34. [doi: 10.3390/geriatrics9020034] [Medline: 38525751]
53. Kim V, Criner GJ. Chronic bronchitis and chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2013 Feb 1;187(3):228-237. [doi: 10.1164/rccm.201210-1843CI] [Medline: 23204254]
54. Oelsner EC, Loehr LR, Henderson AG, et al. Classifying chronic lower respiratory disease events in epidemiologic cohort studies. Ann Am Thorac Soc 2016 Jul;13(7):1057-1066. [doi: 10.1513/AnnalsATS.201601-063OC] [Medline: 27088163]
55. Pesce G. Mortality rates for chronic lower respiratory diseases in Italy from 1979 to 2010: an age-period-cohort analysis. ERJ Open Res 2016 Jan;2(1):00093-2015. [doi: 10.1183/23120541.00093-2015] [Medline: 27730182]
56. Czwikla J, Wandscher K, Helbach J, Fassmer AM, Schmiemann G, Hoffmann F. Prevalence of indwelling urinary catheters in nursing home residents: systematic review. Int J Nurs Stud 2023 Sep;145:104555. [doi: 10.1016/j.ijnurstu.2023.104555] [Medline: 37421830]
57. Czwikla J, Schmiemann G, Hoffmann F. Use of indwelling urinary catheters in nursing home residents: results from a cross-sectional study in 21 German nursing homes. BMC Urol 2024 Jun 14;24(1):125. [doi: 10.1186/s12894-024-01512-w] [Medline: 38877475]
58. Nosocomial infections (healthcare-associated infections). Cleveland Clinic. URL: https://my.clevelandclinic.org/health/diseases/16397-avoiding-healthcare-associated-infections-hais [accessed 2025-09-11]
59. Harrington C, Ross L, Chapman S, Halifax E, Spurlock B, Bakerjian D. Nurse staffing and coronavirus infections in California nursing homes. Policy Polit Nurs Pract 2020 Aug;21(3):174-186. [doi: 10.1177/1527154420938707] [Medline: 32635838]
60. Abrams HR, Loomer L, Gandhi A, Grabowski DC. Characteristics of U.S. nursing homes with COVID-19 cases. J Am Geriatr Soc 2020 Aug;68(8):1653-1656. [doi: 10.1111/jgs.16661] [Medline: 32484912]
61. Chatterjee P, Kelly S, Qi M, Werner RM. Characteristics and quality of US nursing homes reporting cases of coronavirus disease 2019 (COVID-19). JAMA Netw Open 2020 Jul 1;3(7):e2016930. [doi: 10.1001/jamanetworkopen.2020.16930] [Medline: 32725243]
62. Gorges RJ, Konetzka RT. Staffing levels and COVID-19 cases and outbreaks in U.S. nursing homes. J Am Geriatr Soc 2020 Nov;68(11):2462-2466. [doi: 10.1111/jgs.16787] [Medline: 32770832]
63. Sun CLF, Zuccarelli E, Zerhouni EGA, et al. Predicting coronavirus disease 2019 infection risk and related risk drivers in nursing homes: a machine learning approach. J Am Med Dir Assoc 2020 Nov;21(11):1533-1538. [doi: 10.1016/j.jamda.2020.08.030] [Medline: 33032935]
64. White EM, Kosar CM, Feifer RA, et al. Variation in SARS-CoV-2 prevalence in U.S. skilled nursing facilities. J Am Geriatr Soc 2020 Oct;68(10):2167-2173. [doi: 10.1111/jgs.16752] [Medline: 32674223]
65. Zhu X, Lee H, Sang H, et al. Nursing home design and COVID-19: implications for guidelines and regulation. J Am Med Dir Assoc 2022 Feb;23(2):272-279. [doi: 10.1016/j.jamda.2021.12.026] [Medline: 34990585]
66. Tarteret P, Strazzulla A, Rouyer M, et al. Clinical features and medical care factors associated with mortality in French nursing homes during the COVID-19 outbreak. Int J Infect Dis 2021 Mar;104:125-131. [doi: 10.1016/j.ijid.2020.12.004] [Medline: 33301993]
67. Gorges RJ, Konetzka RT. Factors associated with racial differences in deaths among nursing home residents with COVID-19 infection in the US. JAMA Netw Open 2021 Feb 1;4(2):e2037431. [doi: 10.1001/jamanetworkopen.2020.37431] [Medline: 33566110]
68. Gopal R, Han X, Yaraghi N. Compress the curve: a cross-sectional study of variations in COVID-19 infections across California nursing homes. BMJ Open 2021 Jan 6;11(1):e042804. [doi: 10.1136/bmjopen-2020-042804] [Medline: 33408208]
69. Li Y, Temkin-Greener H, Shan G, Cai X. COVID-19 infections and deaths among Connecticut nursing home residents: facility correlates. J Am Geriatr Soc 2020 Sep;68(9):1899-1906. [doi: 10.1111/jgs.16689] [Medline: 32557542]
70. Sugg MM, Spaulding TJ, Lane SJ, et al. Mapping community-level determinants of COVID-19 transmission in nursing homes: a multi-scale approach. Sci Total Environ 2021 Jan 15;752:141946. [doi: 10.1016/j.scitotenv.2020.141946] [Medline: 32889290]
71. Bui DP, See I, Hesse EM, et al. Association between CMS quality ratings and COVID-19 outbreaks in nursing homes - West Virginia, March 17-June 11, 2020. MMWR Morb Mortal Wkly Rep 2020 Sep 18;69(37):1300-1304. [doi: 10.15585/mmwr.mm6937a5] [Medline: 32941409]
72. Cai S, Yan D, Intrator O. COVID-19 cases and death in nursing homes: the role of racial and ethnic composition of facilities and their communities. J Am Med Dir Assoc 2021 Jul;22(7):1345-1351. [doi: 10.1016/j.jamda.2021.05.002] [Medline: 34062147]
73. Konetzka RT. Testimony of R. Tamara Konetzka, PhD. in United States Senate Special Committee on Aging. : United States Senate Special Committee on Aging; 2020 URL: https://www.aging.senate.gov/imo/media/doc/SCA_Konetzka_05_21_20.pdf [accessed 2025-09-11]
74. Chen MK, Chevalier JA, Long EF. Nursing home staff networks and COVID-19. Proc Natl Acad Sci U S A 2021 Jan 7;118(1):e2015455118. [doi: 10.1073/pnas.2015455118] [Medline: 33323526]
75. Cronin CJ, Evans WN. Nursing home quality, COVID-19 deaths, and excess mortality. J Health Econ 2022 Mar;82:102592. [doi: 10.1016/j.jhealeco.2022.102592] [Medline: 35104669]

76. Dean A, Venkataramani A, Kimmel S. Mortality rates from COVID-19 are lower in unionized nursing homes. Health Aff (Millwood) 2020 Nov;39(11):1993-2001. [doi: 10.1377/hlthaff.2020.01011] [Medline: 32910688]
77. Figueroa JF, Wadhera RK, Papanicolas I, et al. Association of nursing home ratings on health inspections, quality of care, and nurse staffing with COVID-19 cases. JAMA 2020 Sep 15;324(11):1103-1105. [doi: 10.1001/jama.2020.14709] [Medline: 32790822]
78. Konetzka RT, White EM, Pralea A, Grabowski DC, Mor V. A systematic review of long-term care facility characteristics associated with COVID-19 outcomes. J Am Geriatr Soc 2021 Oct;69(10):2766-2777. [doi: 10.1111/jgs.17434] [Medline: 34549415]
79. Shen K. Relationship between nursing home COVID-19 outbreaks and staff neighborhood characteristics. PLOS ONE 2022;17(4):e0267377. [doi: 10.1371/journal.pone.0267377] [Medline: 35439279]

Abbreviations

CMS: Centers for Medicare and Medicaid Services
MAR: missing at random
MDS: Minimum Data Set
NHC: nursing home care
PCA: principal component analysis
QM: quality measure
QoC: quality of care
SPO: structure, process, or outcome
UTI: urinary tract infection
ZI: zero inflation
ZINB: zero-inflated negative binomial

Edited by A Kushniruk, E Borycki; submitted 17.Feb.2025; peer-reviewed by F Irshaidat, SJ Lane; revised version received 16.Sep.2025; accepted 23.Sep.2025; published 29.Jan.2026.

Please cite as:
Bharadwaj S, Ali H
Demystifying Quality Metrics and Unveiling the True Measure of Quality of Care in Nursing Homes: Mixed Effects Analysis
JMIR Hum Factors 2026;13:e72770
URL: https://humanfactors.jmir.org/2026/1/e72770
doi:10.2196/72770

Original Paper

Online Tiered Screening for Mental Health Problems Among Refugees in Sweden: Validation Study

Jennifer Meurling[1], MSc; Elisabet Rondung[1], PhD; Youstina Demetry[2], MSc; Anahita Geranmayeh[2], MSc; Anna Leiler[1], PhD; Gerhard Andersson[3], PhD; Anna Bjärtå[1], PhD

[1]Department of Education, Psychology and Social Work, Mid Sweden University, Östersund, Sweden
[2]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden
[3]Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden

Corresponding Author:
Jennifer Meurling, MSc
Department of Education, Psychology and Social Work
Mid Sweden University
Kunskapens väg 1
Östersund, 831 24
Sweden
Phone: 46 101428589
Email: jennifer.meurling@miun.se

Abstract

Background: Refugees and asylum-seekers commonly experience numerous adverse and traumatic events and are therefore at increased risk of mental health problems. Despite the high need for mental health interventions, services tend to be underused by refugees and asylum-seekers, and various barriers compromise access. Digital, efficient screening, adapted for these groups, could facilitate initial assessment and increase accessibility to mental health services. We developed an internet-based tiered screening procedure (i-TAP) aiming to identify clinically relevant symptoms of major depressive disorder (MDD), anxiety disorder, posttraumatic stress disorder, and insomnia disorder among individuals with a refugee background. The i-TAP is an adaptive procedure with 3 tiers aiming to identify general mental distress in Tier 1, differentiate between symptoms in Tier 2, and assess the severity of symptoms in Tier 3. Each tier additionally functions as a gateway to further assessment, as a negative outcome terminates the procedure.

Objective: The purpose of this study was to evaluate the diagnostic test accuracy of the i-TAP, using structured clinical assessments as the reference standard.

Methods: In this prospective study, 70 adult participants with a refugee background, literate in Arabic, Dari, Farsi, or Swedish, and residing in Sweden, completed the i-TAP on tablets and participated in a subsequent structured diagnostic interview.

Results: It has been shown that the i-TAP could identify 91.7% (33/36) of individuals assessed with any psychiatric disorder, and correctly identified 82.1% of all positive cases of MDD, anxiety disorder, posttraumatic stress disorder, and insomnia disorder, with few false negative assessments. Overall test accuracy of the i-TAP ranged between 77.1% and 84.3%, depending on disorder. The tiered design could reduce item burden while maintaining accuracy. A vast majority of participants rated the user experience as positive. In this sample, 36/70 (51.4%) individuals were assessed with one or more psychiatric disorders and comorbidity was high.

Conclusions: The i-TAP may be a valid, efficient, and feasible screening tool for the identification of common psychiatric disorders among individuals with a refugee background in Sweden. The i-TAP could be implemented as a first screener in various settings, including online and in-person clinical practices. The digital, adaptive, multilingual format could facilitate early assessment and increase the availability of mental health services for refugees and asylum-seekers.

(JMIR Hum Factors 2026;13:e82763) doi:10.2196/82763

KEYWORDS

asylum-seekers; diagnostic test accuracy; digital mental health; eHealth; mHealth; online assessment; refugees; tiered screening; validation

Introduction

Background

According to the United Nations High Commissioner for Refugees (UNHCR), there were an estimated 123.2 million forcibly displaced people in the world by the end of 2024 [1]. This number is expected to rise due to ongoing conflicts and wars globally. UNHCR reserves the term refugees for forcibly displaced people who, due to violence, war, conflict, or persecution, have crossed international borders to seek safety in another country, and the term asylum-seeker for people who have applied for refugee status and are awaiting a decision [2]. In Sweden, asylum-seekers have come mainly from Syria, Afghanistan, Iraq, Eritrea, Somalia, and Iran in the years 2015-2025 [3]. Since 2022, refugees from Ukraine have been treated under the European Union Temporary Protection Directive.

In addition to having fled their homes, many refugees and asylum-seekers have endured numerous stressful and potentially traumatic events, increasing the risk of psychological distress and various mental disorders [4-6]. When arriving in their host countries, refugees are further exposed to postmigration stressors, such as language barriers, discrimination, long asylum processes, and economic as well as social challenges, all of which are associated with poor mental health [4,6,7].

Posttraumatic stress disorder (PTSD), major depressive disorder (MDD), and anxiety disorders are highly prevalent among refugees [8,9], with even higher rates observed among asylum-seekers [10,11]. Symptoms persist over time [9,12] and often co-occur [11]. Furthermore, insomnia disorder is highly prevalent [13,14] and various sleep disturbances are reported as problematic [15]. Although prevalence rates vary between studies of refugee mental health, consensus in the field establishes significantly elevated levels of psychiatric symptoms and disorders among refugee groups, including asylum-seekers, compared to the general population [8,9,16]. Despite the high prevalence of mental health issues, mental health services tend to be underused [17,18]. This discrepancy, or treatment gap, can be understood through barriers frequently reported by both help-seeking refugees and asylum-seekers, and service providers. These barriers include language obstacles and the need to use interpreters, difficulties navigating the health care system, and fear of stigma [17,19], along with practical issues and financial concerns [20], all of which affect accessibility to relevant services.

Given that the mental health of individuals who have fled their homelands is affected by adversity experienced before, during, and after the flight [6], the increased risk of mental health problems persists over time, also long after resettlement [9,12]. This underscores the importance of identifying those in need of interventions, regardless of the duration of displacement. In a recent review [21], systematic mental health screening with at least 2 screening occasions is recommended for recently resettled refugees and asylum-seekers. Early and repeated mental health screening could increase access to and provision of treatment, ultimately reducing unnecessary suffering from untreated symptoms. This emphasizes the need for validated, efficient, and feasible methods to identify common mental health problems among individuals who have fled their homelands.

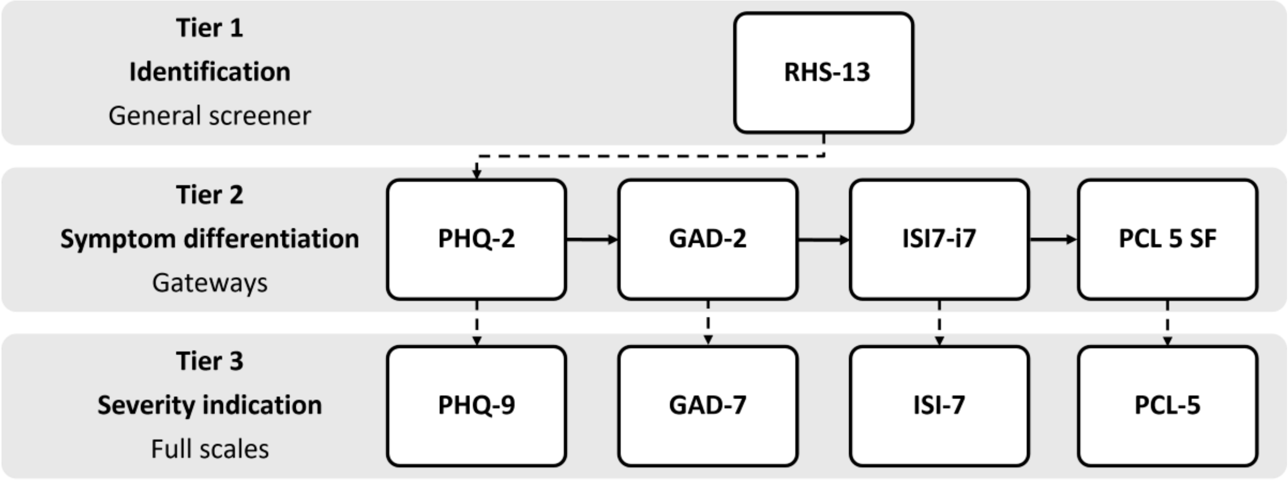
Offering internet-based mental health screening could increase accessibility of mental health services by addressing several of the aforementioned barriers. For instance, internet-based and smartphone-assisted options provide flexibility and privacy, can be time-saving, and can easily be offered in multiple languages [22,23]. Although the literature on this topic is limited, digital mental health screening tools have previously been used among refugees and asylum-seekers [24], with evidence supporting their acceptability, feasibility, and validity for these populations [25,26]. Furthermore, a digital format enables the use of adaptive hierarchical screening models, which allow simultaneous screening of multiple psychiatric symptoms and can reduce the overall respondent burden while maintaining high accuracy [27,28].

Recognizing the potential technical as well as social advantages of online mental health screening for refugees and asylum-seekers, we set out to design and evaluate an internet-based tiered assessment procedure (i-TAP). The i-TAP is tailored for refugee groups residing in Sweden, and specifically designed to identify clinically relevant symptoms of depression, anxiety, PTSD, and insomnia.

The i-TAP

In a previous study by Meurling et al [28], the instruments and the procedure constituting the i-TAP were evaluated, yielding a model designed for optimal psychometric performance, along with promising results regarding the efficiency, including a reduced item burden and accuracy of the procedure. The i-TAP is a 3-tiered screening procedure that adapts to the respondent's answers in each tier. The first tier is highly sensitive, aiming to identify general mental distress while preventing further assessment of individuals without psychiatric symptoms. Thus, a negative outcome in Tier 1 terminates the procedure, while a positive outcome forwards the respondent to Tier 2. The second tier comprises brief symptom scales and differentiates between symptoms of depression, anxiety, PTSD, and insomnia. It also serves a gateway function, forwarding individuals with a positive outcome on one or more scales to further assessment in Tier 3, while terminating the procedure for respondents with a negative outcome on all Tier 2 scales. In the third tier, the full scales corresponding to the outcome of Tier 2 are presented to the respondent, with the purpose of indicating severity and identifying clinically relevant symptoms of depression, anxiety, PTSD, and insomnia. The items already answered in Tier 2 have been removed from each full scale in Tier 3. The tiered design of the i-TAP with the included instruments is depicted in Figure 1. In both the initial study [28] and this one, the i-TAP was provided in Arabic, Dari, Farsi, English, and Swedish, based on the largest groups of refugees and asylum-seekers in Sweden at the time the study commenced [29].

Figure 1. Illustration of the internet-based tiered screening procedure, i-TAP, with the symptom scales used in each tier. GAD-2: Generalized Anxiety Disorder-2; GAD-7: Generalized Anxiety Disorder-7; Insomnia Severity Index-7, item 7; ISI-7: Insomnia Severity Index-7; PCL 5: PTSD (posttraumatic stress disorder) Checklist for DSM-5 (Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]); PCL-5 SF: PTSD (posttraumatic stress disorder) Checklist for DSM-5 (Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]) Short Form; PHQ-2: Patient Health Questionnaire-2; PHQ-9: Patient Health Questionnaire-9; RHS-13: Refugee Health Screener-13.



In Meurling et al [28], the Tier 1 and Tier 2 scales were evaluated using the Tier 3 scales as the reference standard. We also applied receiver operating characteristic analysis to identify gateway cutoffs for optimal performance. This resulted in an accurate screening tool (total precision 86%) that could identify most moderate symptoms of PTSD, depression, anxiety, and/or insomnia (96.2%). These results and the fact that only the first 2 tiers of the model were tested prompted further evaluation of the full 3-tiered model. The developmental process of the i-TAP is described in further detail in Meurling et al [28].

Aim

The aim of this study is to test the performance of the full i-TAP. Specifically, we seek to evaluate the criterion validity by assessing diagnostic test accuracy of the full tiered model, using structured clinical assessments of psychiatric disorders as the reference standard. Our objective is to determine the extent to which the i-TAP is able to identify MDD, anxiety disorder, PTSD, and insomnia disorder among individuals with a refugee background, residing in Sweden. We want to emphasize that the heightened risk for these disorders follows experiences of flight and resettlement. The target group for the i-TAP is therefore focused on the refugee experience, irrespective of current legal status. Accordingly, we henceforward use the term individuals with a refugee background to describe our study population.

Methods

Study Context and Design

This study and the development of the i-TAP are part of the SAHA project, a collaboration between Mid Sweden University, Linköping University, and Karolinska Institute, with the overarching aim of developing digital mental health solutions, including tailored interventions, for individuals with a refugee background. We conducted this prospective cross-sectional validation study with individuals with a refugee background residing in Sweden, collecting data from June to October 2022.

Recruitment and Participants

Participants were recruited through convenience sampling in various nonclinical settings (adult education for immigrants, language cafés, nongovernmental organizations, and at an asylum housing facility). Recruitment took place on site with the help of local personnel and interpreters. After being informed about the study orally, interested individuals could approach immediately or later via email or phone. The inclusion criteria were refugee background, aged 18 years or older, literacy in Arabic, Dari, Farsi, Swedish, or English, and currently living in Sweden. Being a refugee or having a refugee background was herein specified as having fled to Sweden due to war, conflict, persecution, threat, or similar reasons, and was self-defined by the participants. This criterion was carefully explained during recruitment and verified before starting each study procedure, to ensure inclusion of participants with a refugee experience. Individuals were thus included regardless of residence status, and we have not differentiated analyses based on that.

Data Collection and Material

Overview

Data collection was carried out by 11 persons trained in clinical psychology and with practical experience of psychiatric assessment, herein referred to as psychologists. The assessments were conducted in each respondent's preferred language. Five of the psychologists were bilingual and 6 psychologists used authorized interpreters via telephone to conduct the assessments. To facilitate translation, the assessment material was sent to interpreters in advance. All psychologists lived in Sweden and were fluent in Swedish. To ensure methodological equivalence, all psychologists received training in the full procedure prior to data collection.

Data collection included digital screening with the i-TAP, followed by a structured clinical interview. The whole procedure was carried out by a psychologist on a single occasion in a private space adjacent to each recruitment site. Each occasion lasted a maximum of 2 hours and was preceded by oral and written informed consent from the participant.

Digital Screening

Participants completed the i-TAP independently, online via tablets. For this study, we also included background questions and feedback questions regarding the experience of completing the i-TAP. The psychologists (and telephone interpreters) were available to answer any questions but were instructed not to sit next to or help the participants fill in the survey. The results of the i-TAP were blinded to the psychologists.

Structured Diagnostic Interview

The structured diagnostic interview targeted the same disorders as the i-TAP, namely insomnia, PTSD, MDD, and anxiety disorders (panic disorder, agoraphobia, social anxiety disorder, and generalized anxiety disorder [GAD]). Before terminating the clinical interview, the psychologist invited the participant to add any information or raise other topics of their choice. To conclude the procedure, the psychologist summarized the diagnostic interview and gave participants feedback and recommendations based on their assessment. All participants received a brief written summary of the clinical assessment, recommendations (if applicable), a folder with contact information to available local and national health care services, and a SEK 99 (US $9.7) gift card to a supermarket. We offered to guide participants in help-seeking, for example, by showing websites, finding opening hours and contact information, and exploring maps. If an immediate risk was identified, this was handled according to a predefined safety protocol described below. To ensure confidentiality, data were pseudonymized. In cases where the procedure was terminated before completion, participants were also offered the information folder and gift card.

Each clinical interview was discussed at team meetings, referred to as clinical conferences, before determining the final clinical assessment of each participant. The psychologists were still blinded to the results of the i-TAP during these discussions. The clinical conferences promoted both discussion and unanimity in the final assessments.

Index Test

The i-TAP builds on 56 items from 5 validated scales, described in detail below. The symptom scales were selected based on psychometric properties, cross-cultural validity or previous use in refugee populations (or a combination thereof). All scales have previously been evaluated by us for the current population [28,30]. For a detailed description of the scales, translation, and psychometric properties, please see our previous publication from the SAHA project [28]. All scales use Likert response alternatives. For the analyses in this study, responses were dichotomized using cutoffs on each scale. Considering that i-TAP aims for high diagnostic sensitivity yet being an efficient screening procedure, we have applied the sensitivity model proposed in our first study [28] for this study. That is, we have applied cutoffs to identify as many true positives as possible without compromising accuracy.

The following instruments and cutoff values for moderate symptoms were used: for identification of general psychological distress in Tier 1, we used the Refugee Health Screener-13 (RHS-13) [31,32] with cutoff ≥11. For symptoms of MDD, we used the Patient Health Questionnaire-2 (PHQ-2) [33], with cutoff ≥2 in Tier 2 and the Patient Health Questionnaire-9 (PHQ-9) [34] with cutoff ≥10 in Tier 3. Symptoms of anxiety disorders (including panic disorder, social anxiety disorder, and GAD), were measured using the Generalized Anxiety Disorder-2 (GAD-2) [35], with cutoff ≥2 in Tier 2 and the Generalized Anxiety Disorder-7 (GAD-7) [36,37] with cutoff ≥10 in Tier 3. Selection of the anxiety scales was based on very good psychometric properties when screening for GAD, panic disorder, and social anxiety disorder [37], which enables the use of 1 instead of 3 separate anxiety scales, reducing the item burden for the respondents. This aligns with the aim of the i-TAP to efficiently detect clinically relevant symptoms of a range of anxiety disorders. Symptoms of PTSD were measured with the PTSD Checklist for *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [*Fifth Edition*]) Short Form (PCL-5 SF) [38], using cutoff ≥5 in Tier 2 and the PTSD Checklist for *DSM-5* (PCL-5) [39] using cutoff ≥32 in Tier 3. To identify symptoms of insomnia, we used the Insomnia Severity Index-7 (ISI-7) [40], item 7 with cutoff ≥2 in Tier 2 and the full scale in Tier 3. For ISI-7, we previously used a conservative cutoff for detecting moderate to severe symptoms of insomnia (≥14). However, 2 more sensitive cutoffs have been proposed, ≥8 and ≥11 [41,42]. We evaluated the performance of all 3 cutoffs and concluded that cutoff 11 was best suited for the purpose of the i-TAP. Please see Table S1 in Multimedia Appendix 1 for detailed results of this analysis. All Tier 2 instruments are composed of items from the Tier 3 instruments. To avoid posing the same questions twice, these items were not repeated in Tier 3. Instead, we included the item responses from Tier 2 in the analyses of the full scales in Tier 3.

To assess user experience of the i-TAP, we formulated 3 items about answering questions about mental health online. The first assessed if participants had a negative or positive attitude (1-10 with 1=very negative and 10=very positive), the second assessed if the i-TAP was difficult or easy to use (1-10 with 1=very difficult and 10=very easy), and the third assessed if the questions in the i-TAP were difficult or easy to understand (1-5 with 1=very difficult, 2=difficult, 3=neither difficult nor easy, 4=easy, and 5=very easy to understand).

Reference Standard

The reference standard was a clinical diagnosis of MDD, anxiety (panic disorder, agoraphobia, social anxiety disorder, and GAD), PTSD, or insomnia according to the clinical assessment. The assessment of clinical diagnoses relied on a structured diagnostic interview, including sections investigating insomnia disorder, trauma experience, PTSD, MDD, panic disorder, agoraphobia, social anxiety disorder, and GAD, in the presented order. Structured diagnostic interviews are generally accepted as the gold, or reference standard in clinical research [43].

The principal instrument of the structured diagnostic interview was the Mini International Neuropsychiatric Interview 7.0.0 (M.I.N.I.) [44]. M.I.N.I. is a short, structured psychiatric interview based on *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*) and *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) diagnostic criteria and has demonstrated excellent test and interrater reliability [45]. It was originally designed for epidemiological and multicenter studies and is now used extensively in both research and clinical settings, as well as in global mental health settings [46] and with refugee populations [16,47]. The M.I.N.I. is designed to be sensitive [44] and has repeatedly been proven so in comparison with expert clinical diagnoses [43,46]. In this study, we used the diagnostic sections for PTSD, major depressive episode, panic disorder, agoraphobia, social anxiety disorder, and GAD. No official validated translations of the M.I.N.I. were available. For Arabic, we used a previously evaluated translation [46], and the bilingual psychologists made translations to Dari and Farsi. As insomnia is not included in the M.I.N.I., we designed a module for this specific diagnosis comprising 6 questions responding to the diagnostic criteria in *DSM-5* [48], matching the format in the M.I.N.I. to facilitate administration. We formulated brief questions to simplify interpreter-assisted assessments. The complete insomnia assessment is depicted in Multimedia Appendix 1. Furthermore, we adapted the M.I.N.I. to the target population by replacing the suicidality module and the criterion A assessment for PTSD (see details below).

To investigate potentially traumatic events (PTEs), we used the Refugee Trauma History Checklist (RTHC), a measure developed for adult refugees and previously validated in a sample of Syrian refugees in Sweden [49]. The RTHC is designed to assess the most common refugee-related PTEs, before, during, and after flight, in a nonintrusive way. The checklist is composed of 2 questions, where the respondent is asked if they have experienced any of 7 or other PTEs before and after they left their home (war at close quarters, forced separation from family or close friends, loss or disappearance of family members or loved ones, physical violence or assault, witnessing physical violence or assault, torture, sexual violence, and other frightening situation where you felt your life was in danger). Participants were not asked to elaborate or describe reported events further. We used the RTHC as a replacement for item H1 in the M.I.N.I., which corresponds to the criterion A assessment for PTSD in *DSM-5*; thus, a positive RTHC outcome regarding any traumatic life event was followed by continued PTSD assessment with the M.I.N.I. This adaptation was made against previous research pointing toward the importance of broadening the trauma criterion and description in *DSM-5* for individuals with a nonwestern background [50].

To assess suicidality, we used the Columbia Suicide Severity Rating Scale Screening Version (C-SSRS Screen) [51,52] as a structured interview. The C-SSRS Screen has 6 items with a similar structure to the M.I.N.I., is validated as an initial step to identify individuals with short-term suicide risk [51], and has previously been used in refugee populations [53]. A positive outcome was handled according to a predefined safety protocol, including immediate consultation with relevant health services, together with participants, for thorough suicide risk evaluation. If suggested, we offered to accompany the participant to the care facility. We included suicidality as part of the clinical assessment to identify any individuals in need of immediate clinical attention; however, suicidality was not part of this validation study. Against this, and to further adapt the structured assessment to the target group, we decided to replace the extensive and sometimes linguistically complicated suicidality module in the M.I.N.I. with a shorter, more comprehensible screener.

Data Treatment and Analysis

Sample size requirements were estimated based on expected prevalence and desired sensitivity. Based on previous studies of refugee mental health [8,9,16], we expected a high prevalence (>30%) of moderate symptoms of each disorder and an even higher prevalence based on total outcomes from the i-TAP. According to Bujang and Adnan [54], a sample size of 67 is adequate to determine accuracy (80% power and significance level <.05) based on a minimum prevalence of 30%, and the aim of 80% sensitivity we set for the i-TAP.

The structured diagnostic interview was assessed and summarized as positive (“diagnosis”) or negative (“no diagnosis”) for each diagnosis. A diagnosis is herein defined as fulfilling diagnostic criteria for a psychiatric disorder according to clinical assessment with M.I.N.I. and the structured clinical interview for insomnia. Panic disorder, agoraphobia, social anxiety disorder, and GAD in M.I.N.I. were collapsed into one variable, anxiety disorder, for the analysis.

Descriptive measures were summarized as frequencies and proportions or means and SDs depending on variable type. Internal consistency was calculated for the full scales included in the i-TAP using Cronbach alpha. Intercorrelations were assessed using Pearson *r*.

We tested the criterion validity of the i-TAP by comparing the outcome of the complete i-TAP procedure for each symptom category (depression, anxiety, PTSD, and insomnia) with the corresponding clinical diagnoses according to the reference standard. We used Cohen κ as a statistically adjusted measure of concordance between 2 measures [55]. Interpretation of κ values was guided by Landis and Koch [56] suggested labels: <0 no agreement, 0-0.20 slight, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, and 0.81-1.0 perfect agreement. Diagnostic test accuracy, including sensitivity, specificity, positive and negative predictive values [57], and the number of false negatives and false positives, was calculated for each tier and symptom category, using clinical diagnosis from the structured clinical interview as the reference standard.

All individuals had complete data on all scales and were included in the analysis. Analyses were performed using IBM SPSS (version 29).

Ethical Considerations

The study protocol has been approved by the Swedish Ethical Review Authority (2020-00214). This study was planned and performed in accordance with the ethical standards of the 1975 Helsinki declaration and its amendments (World Medical Association, 2001). Written informed consent was obtained from all participants, in their preferred language, prior to study participation. Participants were carefully informed that participation was voluntary and data confidential, that they could terminate their participation at any time, and that participation could not affect the asylum process. To protect the privacy of participants, data are presented at the group level, and potentially identifying details regarding participants or data collection have been omitted.

Results

Sample

In total, 85 individuals registered interest to participate in the study; however, 6 were lost due to nonattendance. Out of the 79 that initiated assessments, 7 did not complete for reasons such as low literacy, not having a refugee background, or practical circumstances. We completed assessments with 72 participants. Due to technical issues, data were incomplete for 2 respondents, rendering a final sample of 70 participants. Against the power calculation, our sample size of 70 was deemed sufficient.

Clinical and Sociodemographic Characteristics

Out of the 70 analyzed assessments, 31 were conducted in Arabic, 14 in Dari, 15 in Farsi, and 10 in Swedish. Participants were 31 females and 39 males, between 19 and 69 years old (mean 39.1, SD 13.16), and 31.4% (22/70) lived at an asylum facility, 50% (35/70) had their own housing, and 18.6% (13/70) lived together with family or friends, see Table 1 for sociodemographic characteristics of the sample.

Table 1. Sample characteristics of participants (N=70).

Characteristic	Participants, n (%)
Nationality	
Syria	25 (35.7)
Afghanistan	19 (27.1)
Iran	12 (17.1)
Palestine	5 (7.1)
Other (n<5)[a]	9 (12.9)
Sex	
Female	31 (44.3)
Male	39 (55.7)
Marital status	
Single	17 (24.3)
Married or partner	44 (62.9)
Divorce or separated	6 (8.6)
Widowed	2 (2.9)
Other	1 (1.4)
Time in Sweden	
<1 year	10 (14.3)
1-3 years	17 (24.3)
4-6 years	19 (27.1)
7-10 years	24 (34.3)
Education	
University Master of Arts	10 (14.3)
University Bachelor of Arts	10 (14.3)
High school	25 (35.7)
Primary school	19 (27.1)
Vocational qualification	3 (4.3)
Other	3 (4.3)
Residence permit	
No residence permit	19 (27.1)
Temporary	18 (25.7)
Permanent	33 (47.1)

[a] Iraq, Eritrea, Tajikistan, Somalia, Sudan, and Kurdistan.

Each scale in the i-TAP demonstrated excellent internal consistency, with Cronbach α of 0.93 for RHS-13, 0.92 for PHQ-9, 0.92 for GAD-7, 0.96 for PCL-5, and 0.92 for ISI-7. All scales were intercorrelated, *r* ranging from 0.70 to 0.89 (all *P*<.001). Regarding symptom assessment, the mean for RHS-13 was 21.61 (SD 13.95; 95% CI 18.29-24.92), for PHQ-9, 9.86 (SD 7.20; 95% CI 8.14-11.57), for GAD-7, 8.33 (SD 6.43; 95% CI 6.79-9.86), for PCL-5, 29.99 (SD 22.74; 95% CI 24.56-35.41), and the mean for ISI-7 was 11.64 (SD 7.96; 95% CI 9.75-13.54).

Prevalence of clinical diagnoses of MDD, anxiety disorder, PTSD, and insomnia disorder was high in the sample, see Table 2 for prevalence estimates for each disorder by the i-TAP and the clinical assessment, respectively. All participants reported experiencing at least one potentially traumatic event. According to the structured diagnostic interview, 51.4% (36/70) met diagnostic criteria for one or more disorders.

Table 2. Prevalence estimates assessed with the structured diagnostic interview and the i-TAP[a], respectively (N=70)[b].

Disorder	Clinical assessment (diagnosis), n (%)	i-TAP (moderate symptoms), n (%)
Major depressive disorder	20 (28.6)	33 (47.1)
Anxiety disorder (any)[c]	17 (24.3)	25 (35.7)
Panic disorder	5 (7.1)	—[d]
Agoraphobia	0 (0)	—
Social anxiety disorder	3 (4.3)	—
Generalized anxiety disorder	14 (20.0)	—
PTSD[e]	20 (28.6)	27 (38.6)
Insomnia disorder	26 (37.1)	32 (45.7)
Any[f]	36 (51.4)	41 (58.6)

[a] i-TAP: internet-based tiered screening procedure.

[b] Multiple diagnoses per individual are possible. Clinical assessment=assessment results of the structured diagnostic interview, i-TAP=individuals who scored positive in Tier 1, Tier 2, and above the cutoff for moderate symptoms in Tier 3.

[c] Positive case for any anxiety disorder.

[d] Not applicable.

[e] PTSD: posttraumatic stress disorder.

[f] Positive case for any disorder.

Comorbidity was generally high. The clinical assessment showed that 23 individuals (32.9%) fulfilled criteria for 2 or more diagnoses, of which 8 individuals met criteria for 3 diagnoses, and 8 for all 4 diagnoses. Comorbidity rates per identified disorder are depicted in Table 3. Among the 13 individuals who fulfilled criteria for one diagnosis only, insomnia was most prevalent (n=7).

Table 3. Comorbidity of disorders diagnosed with the structured diagnostic interview (N=70)[a].

Diagnosis	Number of participants	Comorbid MDD[b], n (%)	Comorbid anxiety disorder, n (%)	Comorbid PTSD[c], n (%)	Comorbid insomnia disorder, n (%)
MDD	20	—[d]	—	—	—
Anxiety disorder	17	12 (17.1)	—	—	—
PTSD	20	14 (20.0)	12 (17.1)	—	—
Insomnia disorder	26	15 (21.4)	13 (18.6)	13 (18.6)	—

[a] Percentages are calculated based on the total sample (N=70).

[b] MDD: major depressive disorder.

[c] PTSD: posttraumatic stress disorder.

[d] Not applicable.

Concordance of the i-TAP and the Structured Diagnostic Interview

The i-TAP identified 33 of the 36 (91.7%) individuals with a clinical diagnosis according to the structured diagnostic interview. The i-TAP correctly identified 19/20 cases of MDD, 13/17 cases of anxiety, 16/20 cases of PTSD, and 21/27 cases of insomnia disorder (ie, true positives) in the current sample. κ values indicated moderate concordance between the i-TAP and the clinical assessment for the specific disorders, and diagnostic test accuracy was between 77.1% and 78.6% for all 4 diagnoses. Regarding overall concordance, accuracy was 84.3% for the identification of any disorder and the overall concordance was substantial, see Table 4 for detailed results.

Table 4. Concordance of the i-TAP[a] and the clinical assessment, calculated for each i-TAP disorder with clinical diagnosis according to the structured diagnostic interview as the reference standard.

i-TAP diagnosis	n_{pos}[b]	Cohen κ	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV[c], % (95% CI)	NPV[d], % (95% CI)	Accuracy, % (95% CI)
Depression	33	.56	95.0 (85.5-100)	72.0 (59.6-84.5)	57.6 (40.7-74.5)	97.3 (92.0-100)	78.6 (68.9-88.2)
Anxiety	25	.46	76.5 (56.3-96,7)	77.4 (66.1-88.7)	52.0 (32.4-71.6)	91.1 (82.8-99.4)	77.1 (67.3-86.9)
PTSD[e]	27	.53	80.0 (62.5-97.5)	78.0 (66.5-89.5)	59.3 (40.8-77.8)	90.7 (82.0-99.3)	78.6 (68.9-88.2)
Insomnia	32	.53	80.8 (64.6-95.4)	75.0 (62.2-87.8)	65.6 (49.1-82.1)	86.8 (76.0-97.6)	77.1 (67.3-86.94)
Any[f]	41	.68	91.7 (82.69-100)	76.5 (62.25-90.75)	80.5 (68.4-92.6)	89.7 (71.1-97.5)	84.3 (75.8-92.8)

[a] i-TAP: internet-based tiered screening procedure.

[b] n_{pos}: number of positive cases for each diagnosis.

[c] PPV: positive predictive value.

[d] NPV: negative predictive value.

[e] PTSD: posttraumatic stress disorder.

[f] Positive case for any diagnosis.

Precision of the i-TAP

To investigate to what level the i-TAP was able to correctly assess the prevalence of disorders on an individual level, we have calculated the number of correct assessments (ie, true positives and true negatives) performed by the i-TAP over all 4 disorders and within each individual. That is, each participant could have 0 to 4 correct assessments of diagnosis or no diagnosis.

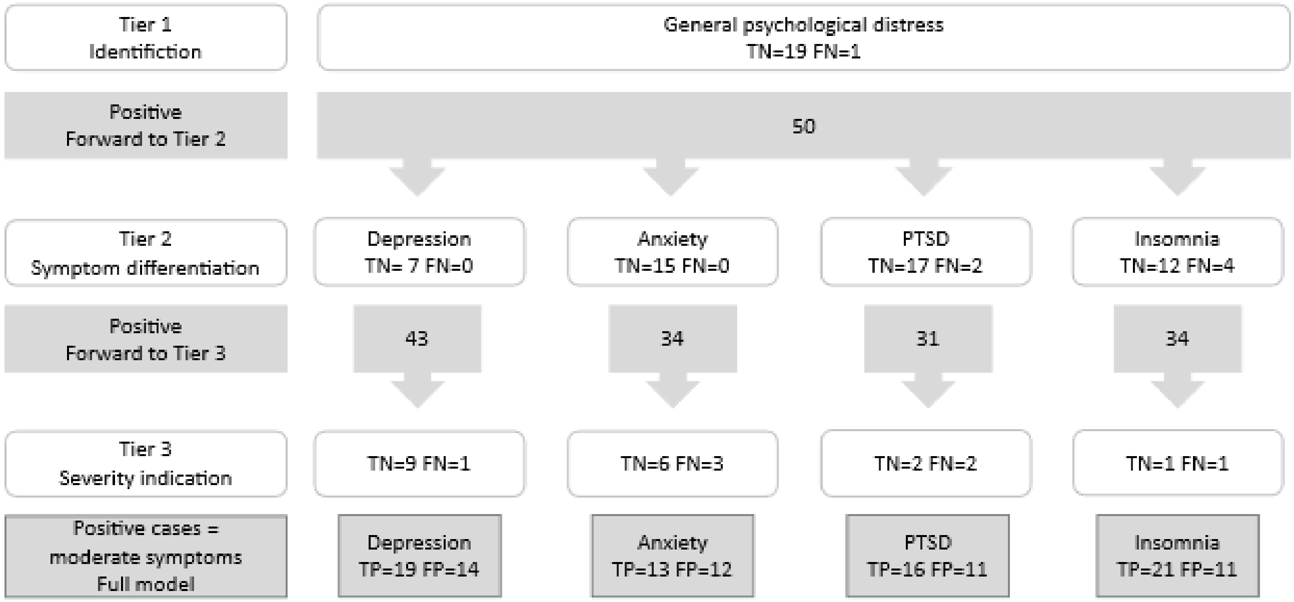
The i-TAP could correctly assess 54.3% (n=38) of all participants. These individuals were thus correctly positively or negatively assessed on all 4 disorders. Fourteen (20.0%) individuals were correctly assessed on 3 of the 4 disorders, of which 11 (78.6%) of the failed assessments were false positives. The remaining 18 participants had 2 or less correctly identified assessments. Equally for these individuals, a vast majority of the failed assessments were false positives (77.1%). A total of 8 individuals were assessed as false positives by the i-TAP without any diagnosis.

Regarding false negative assessments, a total of 11 individuals were affected as the i-TAP failed to identify 1 case of MDD, 4 cases of GAD, 4 cases of PTSD, and 5 cases of insomnia disorder. However, only 3 individuals with a diagnosis according to the reference standard were entirely missed, as the other 8 false negative cases were identified as positive on another diagnosis.

A Tiered Model

To calculate the tiered model, we used all data (all respondents answered all scales) to simulate the effects of the tiers in the i-TAP. In Tier 1, identification of symptoms, 50 individuals screened positive, indicating psychological distress in 71.4% (50/70) of the participants. Tier 1 showed very high sensitivity (97.2%), low specificity (55.9%), and correctly excluded 19 individuals (ie, true negatives) from further assessment; however, 1 individual with clinical anxiety diagnosis was falsely excluded (ie, a false negative). Tier 2, differentiation between symptoms, also exhibited high sensitivity (range 78.6%-100%) and specificity ranging between 54.0% and 74.0% for the 4 disorders. See Figure 2 for an illustration of the culling effects of the hierarchical design. Detailed results of the psychometric evaluation of each tier and disorder are depicted in Table S1 in Multimedia Appendix 2.

Figure 2. The culling effects of the hierarchical design of the internet-based tiered screening procedure (i-TAP). FN: false negatives (incorrectly excluded cases); FP: false positives (incorrectly identified cases); PTSD: posttraumatic stress disorder; TN: true negatives (correctly excluded cases); TP: true positives (correctly identified cases).



User Experience

A majority of the sample had a positive attitude toward answering questions about mental health online, with 51% (36/70) scoring 10/10 (very positive; mean 8.41, SD 2.08; range 3-10) and only 4.3% (3/70) scoring in the negative range (ie, <5). Moreover, the method was perceived as easy (mean 8.47, SD 1.98; range 3-10), with 50% (35/70) scoring 10/10 (very easy), whereas 5.7% (4/70) scored it somewhat difficult (ie, <5). Participants scored the questions as easy to understand, with 61.4% (43/70) scoring 5/5 (very easy; mean 4.37, SD 0.90; range 1-5) and only 1 (1.4%) participant answered that it was very difficult (<3). There were no significant differences between language groups.

Discussion

Principal Findings

The results from this study show that the i-TAP was very efficient in detecting individuals assessed as having a clinical diagnosis (91.7%) and could correctly identify a total of 82.1% of all positive cases of MDD, anxiety disorder, PTSD, and insomnia disorder, while keeping a low rate of false negatives. The overall accuracy of the i-TAP, ranging between 77.1% and 78.6% for the specific diagnoses and rising to 84.3% for any diagnosis, could be considered commendable, considering the multisymptom screening function and the emphasis on sensitivity. Furthermore, the psychometric performance of the i-TAP is comparable to the performance of the individual instruments included [41,50,58,59], showing that a hierarchical design does not compromise accuracy. In line with previous findings [9,15], our results show a sample with high prevalence of MDD, anxiety disorders, PTSD, and insomnia disorder, as well as a high rate of comorbid disorders. The higher prevalence rates found by the i-TAP compared to the clinical assessment were expected, since the use of self-rating instruments is known to overestimate prevalence [60-62] as they identify symptoms, not disorders. Since the i-TAP is a screening tool, we have prioritized sensitivity over specificity in this validation study. This comes at the cost of a higher number of false positives, which have affected the overall diagnostic test accuracy negatively, and which contribute to the overestimation of prevalence. However, prioritizing specificity would increase the risk of overlooking individuals in need. Weighing sensitivity against specificity for the i-TAP is based on the aim to identify as many individuals as possible, and is built upon the standpoint that, in this case, false positives are less harmful than missing someone in need of mental health services. Against this aim, the i-TAP performed very well. However, the relationship between sensitivity and specificity must be considered for each setting and context, with regard to the purpose and resources for implementation of the i-TAP.

Using a tiered hierarchical model for screening enables the reduction of the number of items an individual responds to, referred to as item burden. Reducing the item burden reduces the time and effort required to complete a screening, thereby increasing the probability of both initiating and completing a screening. Using the i-TAP, an individual with no symptoms would have a reduction in item burden with 69.8% compared to a screening using the full scales, and an individual with, for example, clinically relevant symptoms of anxiety would have 37.2% reduction in item burden. However, for the individuals with comorbid symptoms of all disorders, there is, in fact, an increased item burden, adding the RHS-13 [32] in Tier 1. Despite this, we have weighed the benefits of including the RHS-13 as higher than the costs, in this case, in item burden for individuals with symptoms of all 4 diagnoses. The RHS-13 is a suitable first screener, with culturally sensitive items including somatic symptoms, designed for the general refugee population [31]. In this study, it was shown to be highly sensitive, detecting all but one individual and case of clinical diagnosis, thus serving the purpose of the first tier superbly.

The high prevalence of comorbid diagnoses observed in this sample, as evidenced in several prior studies, for example [11,13], underscores the importance of broad initial screening for refugees. Moreover, the high comorbidity might affect the accuracy of the i-TAP for the specific disorders. Several items and symptoms are similar or shared between instruments and diagnoses, making differential diagnosis difficult in early screening. Regarding anxiety specifically, a positive outcome on the i-TAP can imply any, as well as comorbid, anxiety disorders. Furthermore, the i-TAP falsely identified 8 individuals without any disorder. Hence, the i-TAP cannot be used as a diagnostic tool. Nevertheless, the i-TAP could correctly identify a clear majority of the positive cases for each diagnosis, and all but 3 individuals with a clinical diagnosis. Thus, even if not always precise, the i-TAP can be considered a reliable screening tool for the identification of individuals in need of mental health services.

Regarding the user experience of the i-TAP, results were promising, with a vast majority reporting being positive toward the format and finding the questions easy to complete, indicating acceptability of the i-TAP; however, this needs further investigation.

Clinical Implications

The results from this study show that the i-TAP could be implemented as a first screener in various clinical settings, such as primary and psychiatric care, schools, health programs, and online mental health services, for individuals with a refugee background in Sweden. In practice, the i-TAP could be useful as a guide to assessment, for symptom-specific interventions, or to tailor support and treatment. For example, combining the i-TAP with a tailored internet-based cognitive behavioral therapy intervention [63], where treatment module provision is based on symptoms, would facilitate automatic adaptation without requiring help-seeking individuals to answer lengthy questionnaires unnecessarily.

In a setting where differentiation of symptoms is less important, the i-TAP could serve as a general screener to identify individuals with any clinically relevant symptoms, and would do so with higher specificity and efficiency compared to using the RHS-13 or full scales only. Applying either approach, the i-TAP's tiered design could prove especially beneficial in settings with resource and time constraints [25,27]; however, unsuitable in settings where complete data on measures are essential (eg, clinical trials).

Concerning implementation, it is critical to emphasize that the i-TAP is not designed to be a diagnostic instrument and should not be used as such in any context. It is intended as a screening tool, requiring further clinical assessment by a mental health professional in the case of a positive outcome. Moreover, the i-TAP is highly sensitive, and acknowledging the stigma around mental health reported by refugees, we want to highlight the risk of false positive results causing unnecessary distress. Thus, when reporting results back to a patient or participant, it should be done with both cultural and psychometric sensitivity in mind.

Finally, completing the i-TAP requires reading proficiency and digital literacy, and would thus not be a suitable option for all help-seeking individuals with a refugee background; however, adding audio files could be considered to increase applicability. Against this, and the fact that the i-TAP missed 3 individuals with a disorder, we emphasize the need for flexibility and individual adaptation of assessment procedures to promote equity in care.

Limitations

Some methodological limitations need consideration. The selection of participants was neither randomized nor consecutive, resulting in a convenience sample, which raises questions about bias, representativeness, and generalization. Given the sampling method and self-selection, there is a risk that individuals with mental health problems were more likely to participate, potentially resulting in a higher prevalence compared with the broader population of individuals with a refugee background in Sweden. However, relative to this population, our sample included a higher proportion of asylum-seekers (no residence permit) and individuals residing in asylum facilities—established risk factors for mental ill health [10,11]—which likely contributed to the high prevalence found in this study. The literacy requirement excluded illiterate individuals, who thus are not represented in the sample. Nevertheless, regarding the distribution of sex, age, and countries of origin, the sample was found representative of the Swedish adult population of refugees and asylum-seekers at the time of data collection [29]. While the sample mainly comprised individuals from South West Asia, thereby limiting the generalizability of the findings to these groups, it is likely that the i-TAP could be applicable to other refugee populations currently residing in Sweden, such as Ukrainians, as well as in other contexts. However, this warrants future studies.

Another limitation regards language proficiency as we did not control for literacy. However, the results point to a general understanding of the questions, indicating sufficient language proficiency in the sample. Furthermore, about half of the structured diagnostic interviews were conducted with telephone interpreters, the most common interpretation modality within health care in Sweden [64]. Although clinicians, patients, and interpreters have been found to prefer in-person interpretation for building rapport and trust [65,66], a systematic review found no differences in patient satisfaction between modes of interpretation [67]. Validated translations of the M.I.N.I. would have increased reliability and validity, but were not available. This is the common reality in both research and practice with linguistically diverse groups, and interpreters have assisted M.I.N.I. assessments in previous studies [68,69]. Sharing the Swedish and Arabic translations with interpreters did facilitate the interview, but it was not always possible. Nevertheless, the use of interpreters should not, but always can, influence communication quality and ultimately the clinical assessment [9], and thus pose a limitation in this study.

We did not conduct an interrater reliability check as it was not deemed ethically justifiable to film or ask participants to partake twice, given the vulnerability of the population and study settings [70,71]. It would also pose a risk to recruitment. Measures taken to counter this limitation were: psychologists with previous experience of the M.I.N.I., training, authorized interpreters, assessment conferences, and using structured and validated material. We argue that the range of psychologists and interpreters is a strength in this study, reducing the risk of individual bias in assessments.

Using the M.I.N.I. as the reference standard comes with some limitations. Previous studies show that standardized diagnostic interviews result in more diagnoses when compared to clinical psychiatric assessments [43], thus, the results on prevalence should be interpreted with some caution. Regarding the validation of the i-TAP, we however argue that the sensitivity of the M.I.N.I. [43,44] is suitable, as the objective was to develop and evaluate a sensitive screening tool with the aim of identifying clinically relevant symptoms.

Conclusion

Our results suggest that the online tiered assessment procedure, i-TAP, can identify clinical depression, anxiety, PTSD, and insomnia efficiently and with high accuracy among individuals with a refugee background residing in Sweden. Implementation of the i-TAP would enable simultaneous screening of the most common psychiatric disorders among refugees and asylum-seekers, reduce the item burden for symptom-free individuals and circumvent decisions on what to screen for. This, together with the digital, multilingual format of the i-TAP, underlines its potential as a feasible, accessible screening tool, with the potential to bridge several of the reported barriers to mental health care.

Acknowledgments

The authors are grateful to all participants for their time and engagement in the study. We thank Karin Christiansen and her team, as well as Lina Andersson for valuable collaborations and assistance with recruitment. We also acknowledge the contributions of the psychologists Helen Ahmadpourian, Olivia Erling, Tove Forsström, Hania Kutabi, Agnes Lorenz, Hajar Mohammadi, and Thilda Persson. For translations of the material, we thank Amjad AbuLeil, Farzaneh Sianpour, and Sohrab Soltani. Finally, we thank George Vlaescu for providing technical support.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to continued analyses and reports but are available from the corresponding author on reasonable request.

Funding

This study was funded by the Swedish Research Council (2018-05827). The funders did not partake in or influence the data collection, analysis, or reporting of the findings. Open access funding provided by Mid Sweden University.

Authors' Contributions

Conceptualization: AB, JM
Methodology: AB, JM, ER, AL
Investigation: AB, JM, ER, YD, AG
Formal analysis: JM
Project administration: JM
Resources: JM
Supervision: AB, ER
Funding acquisition: GA
Writing – original draft: AB, JM, ER
Writing – review & editing: AB, JM, ER, YD, AG, AL, GA

Conflicts of Interest

None declared.

Multimedia Appendix 1
Assessment material for insomnia disorder and detailed psychometric evaluation of the Insomnia Severity Index-7.
[DOCX File , 40 KB - humanfactors_v13i1e82763_app1.docx]

Multimedia Appendix 2
Extended psychometric evaluation of the internet-based tiered screening procedure.
[DOCX File , 38 KB - humanfactors_v13i1e82763_app2.docx]

References

1. Global trends report - forced displacement in 2024. United Nations High Commissioner for Refugees. 2025. URL: https://wwwunhcrorg/global-trends-report-2024 [accessed 2026-01-07]

2. Convention and protocol relating to the status of refugees. United Nations High Commissioner for Refugees. 2025. URL: https://www.unhcr.org/3b66c2aa102010 [accessed 2026-01-07]
3. Asylsökande under året efter medborgarskapsland och kön [Asylum-seekers in the year, by country of citizenship and sex]. SCB. 2025. URL: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__BE__BE0101__BE0101P/AsylsokandeN/: [accessed 2026-01-07]
4. Bogic M, Ajdukovic D, Bremner S, Franciskovic T, Galeazzi GM, Kucukalic A, et al. Factors associated with mental disorders in long-settled war refugees: refugees from the former Yugoslavia in Germany, Italy and the UK. Br J Psychiatry 2012;200(3):216-223 [FREE Full text] [doi: 10.1192/bjp.bp.110.084764] [Medline: 22282430]
5. Turrini G, Purgato M, Ballette F, Nosè M, Ostuzzi G, Barbui C. Common mental disorders in asylum seekers and refugees: umbrella review of prevalence and intervention studies. Int J Ment Health Syst 2017;11:51 [FREE Full text] [doi: 10.1186/s13033-017-0156-0] [Medline: 28855963]
6. Porter M, Haslam N. Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: a meta-analysis. JAMA 2005;294(5):602-612. [doi: 10.1001/jama.294.5.602] [Medline: 16077055]
7. Malm A, Tinghög P, Narusyte J, Saboonchi F. The refugee post-migration stress scale (RPMS) - development and validation among refugees from Syria recently resettled in Sweden. Confl Health 2020;14:2 [FREE Full text] [doi: 10.1186/s13031-019-0246-5] [Medline: 31921332]
8. Fazel M, Wheeler J, Danesh J. Prevalence of serious mental disorder in 7000 refugees resettled in western countries: a systematic review. Lancet 2005;365(9467):1309-1314. [doi: 10.1016/S0140-6736(05)61027-6] [Medline: 15823380]
9. Blackmore R, Boyle JA, Fazel M, Ranasinha S, Gray KM, Fitzgerald G, et al. The prevalence of mental illness in refugees and asylum seekers: a systematic review and meta-analysis. PLoS Med 2020;17(9):e1003337 [FREE Full text] [doi: 10.1371/journal.pmed.1003337] [Medline: 32956381]
10. Posselt M, McIntyre H, Ngcanga M, Lines T, Procter N. The mental health status of asylum seekers in middle- to high-income countries: a synthesis of current global evidence. Br Med Bull 2020;134(1):4-20. [doi: 10.1093/bmb/ldaa010] [Medline: 32409820]
11. Leiler A, Bjärtå A, Ekdahl J, Wasteson E. Mental health and quality of life among asylum seekers and refugees living in refugee housing facilities in Sweden. Soc Psychiatry Psychiatr Epidemiol 2019;54(5):543-551. [doi: 10.1007/s00127-018-1651-6] [Medline: 30580381]
12. Borho A, Viazminsky A, Morawa E, Schmitt GM, Georgiadou E, Erim Y. The prevalence and risk factors for mental distress among Syrian refugees in Germany: a register-based follow-up study. BMC Psychiatry 2020;20(1):362. [doi: 10.1186/s12888-020-02746-2] [Medline: 32641014]
13. Al-Smadi AM, Tawalbeh LI, Gammoh OS, Ashour A, Tayfur M, Attarian H. The prevalence and the predictors of insomnia among refugees. J Health Psychol 2019;24(8):1125-1133. [doi: 10.1177/1359105316687631] [Medline: 28810381]
14. Basishvili T, Eliozishvili M, Maisuradze L, Lortkipanidze N, Nachkebia N, Oniani T, et al. Insomnia in a displaced population is related to war-associated remembered stress. Stress Health 2012;28(3):186-192. [doi: 10.1002/smi.1421] [Medline: 22282401]
15. Richter K, Baumgärtner L, Niklewski G, Peter L, Köck M, Kellner S, et al. Sleep disorders in migrants and refugees: a systematic review with implications for personalized medical approach. EPMA J 2020;11(2):251-260 [FREE Full text] [doi: 10.1007/s13167-020-00205-2] [Medline: 32549917]
16. Patanè M, Ghane S, Karyotaki E, Cuijpers P, Schoonmade L, Tarsitani L, et al. Prevalence of mental disorders in refugees and asylum seekers: a systematic review and meta-analysis. Glob Ment Health (Camb) 2022;9:250-263 [FREE Full text] [doi: 10.1017/gmh.2022.29] [Medline: 36618716]
17. Satinsky E, Fuhr DC, Woodward A, Sondorp E, Roberts B. Mental health care utilisation and access among refugees and asylum seekers in Europe: a systematic review. Health Policy 2019;123(9):851-863 [FREE Full text] [doi: 10.1016/j.healthpol.2019.02.007] [Medline: 30850148]
18. Due C, Green E, Ziersch A. Psychological trauma and access to primary healthcare for people from refugee and asylum-seeker backgrounds: a mixed methods systematic review. Int J Ment Health Syst 2020;14(1):71 [FREE Full text] [doi: 10.1186/s13033-020-00404-4] [Medline: 32944067]
19. Kiselev N, Morina N, Schick M, Watzke B, Schnyder U, Pfaltz MC. Barriers to access to outpatient mental health care for refugees and asylum seekers in Switzerland: the therapist's view. BMC Psychiatry 2020;20(1):378 [FREE Full text] [doi: 10.1186/s12888-020-02783-x] [Medline: 32680485]
20. Byrow Y, Pajak R, Specker P, Nickerson A. Perceptions of mental health and perceived barriers to mental health help-seeking amongst refugees: a systematic review. Clin Psychol Rev 2020;75:101812. [doi: 10.1016/j.cpr.2019.101812] [Medline: 31901882]
21. Maillet L, Goudet A, Godbout I, Ntanda GM, Laliberté G, Desjardins F, et al. Measuring the optimal time interval between arrival and first mental health evaluation's for refugees in Québec: a scoping review. Adm Policy Ment Health 2023;50(4):563-575 [FREE Full text] [doi: 10.1007/s10488-023-01257-y] [Medline: 36881163]
22. Whitehead L, Talevski J, Fatehi F, Beauchamp A. Barriers to and facilitators of digital health among culturally and linguistically diverse populations: qualitative systematic review. J Med Internet Res 2023;25:e42719 [FREE Full text] [doi: 10.2196/42719] [Medline: 36853742]

23. Liem A, Natari RB, Jimmy, Hall BJ. Digital health applications in mental health care for immigrants and refugees: a rapid review. Telemed J E Health 2021;27(1):3-16. [doi: 10.1089/tmj.2020.0012] [Medline: 32498658]
24. Donnelly O, Leavey G. Screening tools for mental disorders among female refugees: a systematic review. J Child Adolesc Trauma 2022;15(2):209-219 [FREE Full text] [doi: 10.1007/s40653-021-00375-9] [Medline: 35600514]
25. Morina N, Ewers SM, Passardi S, Schnyder U, Knaevelsrud C, Müller J, et al. Mental health assessments in refugees and asylum seekers: evaluation of a tablet-assisted screening software. Confl Health 2017;11:18 [FREE Full text] [doi: 10.1186/s13031-017-0120-2] [Medline: 29026438]
26. Vindbjerg E, Sonne C, Silove D, Bibby H, Hall J, Momartin S, et al. Evaluating the procedural validity of the multilingual computer assisted self interview (MultiCASI) in a refugee population. Psychiatry Res 2022;312:114529. [doi: 10.1016/j.psychres.2022.114529] [Medline: 35398659]
27. Batterham PJ, Calear AL, Sunderland M, Carragher N, Christensen H, Mackinnon AJ. Hierarchical screening for multiple mental disorders. J Affect Disord 2013;151(1):229-236. [doi: 10.1016/j.jad.2013.05.085] [Medline: 23806587]
28. Meurling J, Rondung E, Leiler A, Wasteson E, Andersson G, Richards D, et al. An online tiered screening procedure to identify mental health problems among refugees. BMC Psychiatry 2023;23(1):7 [FREE Full text] [doi: 10.1186/s12888-022-04481-2] [Medline: 36597066]
29. Statistik om asylärenden. Migrationsverket. 2022. URL: https://www.migrationsverket.se/Om-Migrationsverket/Statistik.html2022 [accessed 2026-01-07]
30. Bjärtå A, Leiler A, Ekdahl J, Wasteson E. Assessing severity of psychological distress among refugees with the Refugee Health Screener, 13-item version. J Nerv Ment Dis 2018;206(11):834-839 [FREE Full text] [doi: 10.1097/NMD.0000000000000886] [Medline: 30256332]
31. Hollifield M, Toolson E, Verbillis-Kolp S, Farmer B, Yamazaki J, Woldehaimanot T, et al. Effective screening for emotional distress in refugees: the Refugee Health Screener. J Nerv Ment Dis 2016;204(4):247-253. [doi: 10.1097/NMD.0000000000000469] [Medline: 26825376]
32. Hollifield M, Verbillis-Kolp S, Farmer B, Toolson EC, Woldehaimanot T, Yamazaki J, et al. The Refugee Health Screener-15 (RHS-15): development and validation of an instrument for anxiety, depression, and PTSD in refugees. Gen Hosp Psychiatry 2013;35(2):202-209. [doi: 10.1016/j.genhosppsych.2012.12.002] [Medline: 23347455]
33. Kroenke K, Spitzer R, Williams J. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
34. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
35. Sapra A, Bhandari P, Sharma S, Chanpura T, Lopp L. Using Generalized Anxiety Disorder-2 (GAD-2) and GAD-7 in a primary care setting. Cureus 2020;12(5):e8224 [FREE Full text] [doi: 10.7759/cureus.8224] [Medline: 32582485]
36. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
37. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
38. Zuromski KL, Ustun B, Hwang I, Keane TM, Marx BP, Stein MB, et al. Developing an optimal short-form of the PTSD checklist for DSM-5 (PCL-5). Depress Anxiety 2019;36(9):790-800 [FREE Full text] [doi: 10.1002/da.22942] [Medline: 31356709]
39. Weathers FW, Keane TM, Palmieri PA, Marx BP, Schnurr PP. The PTSD Checklist for DSM-5 (PCL-5). 2013. URL: https://www.ptsd.va.gov/professional/assessment/adult-sr/
40. Bastien C, Vallières A, Morin C. Validation of the insomnia severity index as an outcome measure for insomnia research. Sleep Med 2001;2(4):297-307. [doi: 10.1016/s1389-9457(00)00065-4] [Medline: 11438246]
41. Morin C, Belleville G, Bélanger L, Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep 2011;34(5):601-608 [FREE Full text] [doi: 10.1093/sleep/34.5.601] [Medline: 21532953]
42. Morin CM, Colecchi C, Stone J, Sood R, Brink D. Behavioral and pharmacological therapies for late-life insomnia: a randomized controlled trial. JAMA 1999;281(11):991-999. [doi: 10.1001/jama.281.11.991] [Medline: 10086433]
43. Rettew DC, Lynch AD, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. Int J Methods Psychiatr Res 2009;18(3):169-184 [FREE Full text] [doi: 10.1002/mpr.289] [Medline: 19701924]
44. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]
45. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Sheehan KH, et al. The mini international neuropsychiatric interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatr 2020;12(5):224-231. [doi: 10.1016/s0924-9338(97)83296-8]

46. Karnouk C, Böge K, Lindheimer N, Churbaji D, Abdelmagid S, Mohamad S, et al. Development of a culturally sensitive Arabic version of the Mini International Neuropsychiatric Interview (M.I.N.I.-AR) and validation of the depression module. Int J Ment Health Syst 2021;15(1):24 [FREE Full text] [doi: 10.1186/s13033-021-00447-1] [Medline: 33736659]
47. Wiechers M, Strupf M, Bajbouj M, Böge K, Karnouk C, Goerigk S, et al. Empowerment group therapy for refugees with affective disorders: results of a multicenter randomized controlled trial. Eur Psychiatry 2023;66(1):e64 [FREE Full text] [doi: 10.1192/j.eurpsy.2023.2431] [Medline: 37458215]
48. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). Washington: American Psychiatric Association; 2013.
49. Sigvardsdotter E, Nilsson H, Malm A, Tinghög P, Gottvall M, Vaez M, et al. Development and preliminary validation of Refugee Trauma History Checklist (RTHC)-a brief checklist for survey studies. Int J Environ Res Public Health 2017;14(10):1175 [FREE Full text] [doi: 10.3390/ijerph14101175] [Medline: 28976937]
50. Verhey R, Chibanda D, Gibson L, Brakarsh J, Seedat S. Validation of the posttraumatic stress disorder checklist - 5 (PCL-5) in a primary care population with high HIV prevalence in Zimbabwe. BMC Psychiatry 2018;18(1):109 [FREE Full text] [doi: 10.1186/s12888-018-1688-9] [Medline: 29685117]
51. Bjureberg J, Dahlin M, Carlborg A, Edberg H, Haglund A, Runeson B. Columbia-suicide severity rating scale screen version: initial screening for suicide risk in a psychiatric emergency department. Psychol Med 2021;52(16):1-9 [FREE Full text] [doi: 10.1017/S0033291721000751] [Medline: 33766155]
52. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
53. Rondung E, Leiler A, Sarkadi A, Bjärtå A, Lampa E, Löfving SG, et al. Feasibility of a randomised trial of teaching recovery techniques (TRT) with refugee youth: results from a pilot of the Swedish UnaccomPanied yOuth Refugee Trial (SUPpORT). Pilot Feasibility Stud 2022;8(1):40 [FREE Full text] [doi: 10.1186/s40814-022-00998-1] [Medline: 35164865]
54. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. J Clin Diagn Res 2016;10(10):YE01-YE06 [FREE Full text] [doi: 10.7860/JCDR/2016/18129.8744] [Medline: 27891446]
55. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960 Apr 01;20(1):37-46. [doi: 10.1177/001316446002000104]
56. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-174. [doi: 10.2307/2529310]
57. Mandrekar JN. Simple statistical measures for diagnostic accuracy assessment. J Thorac Oncol 2010;5(6):763-764 [FREE Full text] [doi: 10.1097/JTO.0b013e3181dab122] [Medline: 20502268]
58. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
59. Urtasun M, Daray FM, Teti GL, Coppolillo F, Herlax G, Saba G, et al. Validation and calibration of the Patient Health Questionnaire (PHQ-9) in Argentina. BMC Psychiatry 2019;19(1):291 [FREE Full text] [doi: 10.1186/s12888-019-2262-9] [Medline: 31533674]
60. Turner SW, Bowie C, Dunn G, Shapo L, Yule W. Mental health of Kosovan Albanian refugees in the UK. Br J Psychiatry 2018;182(5):444-448. [doi: 10.1192/bjp.182.5.444]
61. Engelhard IM, van den Hout MA, Weerts J, Arntz A, Hox JJCM, McNally RJ. Deployment-related stress and trauma in Dutch soldiers returning from Iraq. Prospective study. Br J Psychiatry 2007;191:140-145. [doi: 10.1192/bjp.bp.106.034884] [Medline: 17666498]
62. Thombs BD, Kwakkenbos L, Levis AW, Benedetti A. Addressing overestimation of the prevalence of depression based on self-report screening questionnaires. CMAJ 2018;190(2):E44-E49 [FREE Full text] [doi: 10.1503/cmaj.170691] [Medline: 29335262]
63. Lindegaard T, Seaton F, Halaj A, Berg M, Kashoush F, Barchini R, et al. Internet-based cognitive behavioural therapy for depression and anxiety among Arabic-speaking individuals in Sweden: a pilot randomized controlled trial. Cogn Behav Ther 2021;50(1):47-66 [FREE Full text] [doi: 10.1080/16506073.2020.1771414] [Medline: 32603632]
64. Tolkutredningen 2019. Almega-Språkföretagen. 2019. URL: https://www.regeringen.se/contentassets/626ccb432337479d9c5968efebb5186a/almega-sprakforetagen.pdf [accessed 2026-01-07]
65. Ruiz A, Chen J, Brown TT, Cai X, Hernandez Fernandez P, Rodriguez HP. Comparing remote and in-person interpretation experiences for clinicians and Spanish-speaking patients with limited English proficiency: a mixed methods study. BMJ Open Qual 2025;14(2):e003227 [FREE Full text] [doi: 10.1136/bmjoq-2024-003227] [Medline: 40340979]
66. Price EL, Pérez-Stable EJ, Nickleach D, López M, Karliner LS. Interpreter perspectives of in-person, telephonic, and videoconferencing medical interpretation in clinical encounters. Patient Educ Couns 2012;87(2):226-232 [FREE Full text] [doi: 10.1016/j.pec.2011.08.006] [Medline: 21930360]
67. Joseph C, Garruba M, Melder A. Patient satisfaction of telephone or video interpreter services compared with in-person services: a systematic review. Aust Health Rev 2018;42(2):168-177. [doi: 10.1071/AH16195] [Medline: 30021688]

68. Heeren M, Mueller J, Ehlert U, Schnyder U, Copiery N, Maier T. Mental health of asylum seekers: a cross-sectional study of psychiatric disorders. BMC Psychiatry 2012;12:114 [FREE Full text] [doi: 10.1186/1471-244X-12-114] [Medline: 22900706]
69. Hocking DC, Mancuso SG, Sundram S. Development and validation of a mental health screening tool for asylum-seekers and refugees: the STAR-MH. BMC Psychiatry 2018;18(1):69 [FREE Full text] [doi: 10.1186/s12888-018-1660-8] [Medline: 29548315]
70. Deps P, Rezende I, Andrade M, Collin S. Ethical issues in research with refugees. Ethics Med Public Health 2022;24:100813. [doi: 10.1016/j.jemep.2022.100813]
71. International ethical guidelines for health-related research involving humans. Council for International Organizations of Medical Sciences WHO. 2016. URL: https://cioms.ch/publications/product/international-ethical-guidelines-for-health-related-research-involving-humans/ [accessed 2026-01-07]

Abbreviations

C-SSRS Screen: Columbia Suicide Severity Rating Scale Screening Version
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
GAD: generalized anxiety disorder
GAD-2: Generalized Anxiety Disorder-2
GAD-7: Generalized Anxiety Disorder -7
ICD-10: International Statistical Classification of Diseases, Tenth Revision
ISI-7: Insomnia Severity Index
i-TAP: internet-based tiered assessment procedure
MDD: major depressive disorder
M.I.N.I.: Mini International Neuropsychiatric Interview
PCL-5: PTSD (posttraumatic stress disorder) Checklist for DSM-5 (Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition])
PCL-5 SF: PTSD (posttraumatic stress disorder) Checklist for DSM-5 (Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]) Short Form
PHQ-2: Patient Health Questionnaire-2
PHQ-9: Patient Health Questionnaire-9
PTE: potentially traumatic event
PTSD: posttraumatic stress disorder
RHS-13: Refugee Health Screener -13
RTHC: Refugee Trauma History Checklist
UNHCR: United Nations High Commissioner for Refugees

Edited by A Kushniruk; submitted 21.Aug.2025; peer-reviewed by DC Hocking, N Kiselev; comments to author 01.Dec.2025; accepted 02.Jan.2026; published 29.Jan.2026.

Please cite as:
Meurling J, Rondung E, Demetry Y, Geranmayeh A, Leiler A, Andersson G, Bjärtå A
Online Tiered Screening for Mental Health Problems Among Refugees in Sweden: Validation Study
JMIR Hum Factors 2026;13:e82763
URL: https://humanfactors.jmir.org/2026/1/e82763
doi:10.2196/82763
PMID:

Objective Monitoring of Tablet Use–Related Optical Exposure and Its Association With Axial Length in Preschool Children: Cross-Sectional Intelligent Monitoring Study

Yidong Zhu[1*], MPH; Hao Chen[1*], PhD; Senlin Lin[2,3*], MSc; Hong Jiang[4*], PhD; Mingdao Zhang[5], MEng; Yi Sun[1], MPH; Chenshu Li[1], MPH; Yingnan Jia[1], PhD

[1]Preventive Medicine and Health Education Department, School of Public Health, Fudan University, 130 Dong An Road, Shanghai, China
[2]Shanghai Eye Diseases Prevention & Treatment Center, Shanghai, China
[3]Shanghai Eye Hospital, School of Medicine, Tongji University, Shanghai, China
[4]Department of Maternal, Child and Adolescent Health, School of Public Health, Fudan University, Shanghai, China
[5]Midea Group (China), Nanchang, Jiangxi, China
[*]these authors contributed equally

Corresponding Author:
Yingnan Jia, PhD
Preventive Medicine and Health Education Department, School of Public Health, Fudan University, 130 Dong An Road, Shanghai, China

Abstract

Background: In recent years, the global prevalence of myopia among children has continued to rise. The preschool years represent a critical period for visual development, and the widespread adoption of electronic screens among young children has brought increasing attention to pediatric visual health. However, the association between visual environmental exposures related to screen use—such as screen brightness and ambient illuminance—and the risk of myopia in preschool children has not been thoroughly investigated.

Objective: This monitoring study aimed to investigate the association between electronic screen brightness, ambient illuminance, and axial length in preschool children to provide evidence-based support for developing screen brightness usage recommendations for this population.

Methods: This cross-sectional monitoring study was conducted between March and July 2023 in Shanghai, China, involving 2 representative samples of kindergarten children aged 3 to 6 years. Each participant was provided with a tablet preinstalled with intelligent monitoring software, which continuously and objectively recorded real-time data on screen time and screen brightness over a consecutive 7-day period. In addition, comprehensive data collection encompassed standardized ophthalmic assessments, high-precision ambient illuminance measurements, simulated laboratory lighting evaluations, and parental questionnaires. Associations between ambient illuminance, screen brightness, and axial length were analyzed using multivariable linear regression and restricted cubic spline models.

Results: Of the 199 children included in the total sample, 124 (62.3%) were boys, and 75 (37.7%) were girls. After adjustment for demographic characteristics, parental myopia, and screen use behaviors, the median ambient illuminance during tablet use was significantly inversely associated with axial length (β=–0.13, 95% CI –0.22 to –0.04; P=.006). A nonlinear dose-response relationship was identified between median screen brightness and axial length ($P_{\text{nonlinearity}}$=.004), with axial elongation accelerating beyond approximately 27 cd/m² and peaking around 56 cd/m². Boys (P<.001) and greater height (P=.33) were also significantly associated with longer axial length.

Conclusions: Higher ambient illuminance during tablet use is associated with shorter axial length in preschoolers, whereas screen brightness exhibits a nonlinear effect on axial elongation. This study highlights the importance of optimizing both environmental lighting and device settings to protect visual health in young children, providing empirical support for guidelines on safe digital device use and ambient lighting conditions in early childhood.

(JMIR Hum Factors 2026;13:e79266) doi:10.2196/79266

KEYWORDS

screen brightness; ambient illuminance; axial length; preschool children; visual health

Introduction

The rising global prevalence of myopia among children and adolescents has become a significant public health challenge [1]. Early intervention is critical, as a younger age of onset is strongly associated with a higher risk of developing high myopia [2,3]. Axial length represents a key biomarker for evaluating myopia severity and predicting disease progression, with abnormal elongation being strongly associated with myopia advancement [4]. The preschool stage (age 3 - 5 y) is a critical period for refractive development. During this time, the characteristics of visual environmental exposure can profoundly influence the onset and progression of myopia.

In our digitalized society, electronic screens have become a dominant component of children’s visual environment. A large-scale epidemiological study conducted across 38 countries on 6 continents revealed that 60% to 93% of children and adolescents are exposed to screens for more than 2 hours daily [5], with preschool children showing a trend of earlier initial screen exposure and higher usage frequency [6,7]. This underscores the urgent need to evaluate the potential impact of screen exposure on visual health and myopia risk in preschoolers. Population-based studies have preliminarily demonstrated the association between screen time and myopia. A systematic review and meta-analysis published in *The Lancet Digital Health* confirmed that the independent duration of smartphone and tablet use as well as the cumulative effect of screen time from computers were significantly associated with an increased risk of myopia [8]. A study from Japan further identified a relationship between prolonged screen time and abnormal axial length elongation in young children [9]. A cross-sectional study from the Netherlands revealed a significant positive correlation between the increased frequency of uninterrupted smartphone usage lasting over 20 minutes and increased scleral elasticity ratio as well as the anterior chamber depth-to-lens thickness ratio [10]. Although existing studies have investigated the association between screen time and visual development, research on the impact of screen brightness, a critical optical parameter, on visual health remains absent and warrants further exploration.

Modern mobile devices are commonly equipped with ambient light sensors that enable automatic screen brightness adjustment, a dynamic feature that may influence the visual system through two mechanisms: first, the adaptation of screen brightness to ambient light levels may affect visual function; and second, excessively high ambient light levels may induce dry eye symptoms, while abnormally elevated screen brightness could increase the burden of glare perception [11-13]. Evidence from animal models indicates that low illumination levels can induce myopia, whereas increased light exposure may activate retinal dopaminergic activity (primarily via the D2 receptor pathway), thereby regulating retinal gene expression and suppressing the biological signaling pathways that promote axial elongation [14]. Furthermore, ergonomic research has identified an inverted U-shaped relationship between display brightness and visual performance, fatigue, and discomfort, suggesting that both excessively low and high brightness levels compromise visual comfort and performance [15]. Benedetto et al [11] demonstrated that increased screen brightness significantly exacerbates visual fatigue. However, epidemiological evidence regarding the impact of screen brightness on eye axial elongation in children remains lacking.

Therefore, the primary aim of this study was to collect real-time objective exposure data through intelligent monitoring technology and then examine the association between electronic screen brightness, ambient illuminance, and axial length in preschool children. These findings may contribute to the development of evidence-based guidelines regarding screen use and myopia prevention in young children.

Methods

Study Design and Population

This cross-sectional monitoring study was conducted in two districts of Shanghai, China: Xuhui District (urban) and Pudong New Area (suburban), from March to July 2023. A random sampling method was used to select a kindergarten from each of the 2 districts. For each kindergarten, a cluster sampling method was used to select 5 classes, in which the average number of children in each class was 30, from the junior, middle, and senior grades. A total of 199 preschool children participated in the study, each of whom was provided with a tablet for a 1-week period. Before participation, parental consent was obtained during the eligibility screening phase. During the monitoring period, participants were instructed to maintain their typical screen usage habits, with the provided tablets being strictly limited to use by child participants only, thereby ensuring the integrity of exposure data collection. The study enrollment criteria were established as follows. Inclusion criteria were as follows: (1) preschool children currently registered in the kindergarten; (2) aged 3 to 6 years at the time of recruitment; and (3) whose parents or other caregivers have allowed them to use screen devices, such as televisions, mobile phones, tablets, and computers. Exclusion criteria were as follows: (1) presence of diagnosed organic ocular pathologies, (2) demonstrated inability to complete refractive error assessments, and (3) absence of legally authorized guardian consent.

Measures

Vision Screening

Axial length, corneal curvature, and refractive errors were measured using standardized ophthalmic instruments. Spherical and cylindrical refractive errors were assessed with the KR-800 auto-refractor (Topcon Corporation), while axial length and corneal curvature were measured using the IOL Master 700 (Carl Zeiss Meditec AG). To ensure measurement accuracy, examinations were conducted sequentially, starting with the right eye, followed by the left eye. If children demonstrated poor cooperation, measurements were repeated. All examiners received prior training, and instruments were regularly calibrated to maintain precision.

Screen Exposure Monitoring Technology

The *Healthy Screen Viewing for Children (HSVC)* application, developed in collaboration with Huawei Corporation, was installed on dedicated tablets provided to participants [16]. This software automatically recorded screen usage data, including application type, screen brightness, and duration of use, at 1-minute intervals over a 7-day monitoring period. Equipped with built-in ambient light sensors, the tablets dynamically adjusted screen brightness according to real-time ambient lighting conditions. The HSVC application concurrently recorded the corresponding screen luminance values (in cd/m²) along with precise time stamps for each measurement. All collected data were securely transmitted to a cloud server for subsequent analysis.

Simulated Lighting Environment Setup

To convert the recorded screen brightness values into corresponding ambient illuminance levels for each child, we established a response model between screen brightness and horizontal illuminance under controlled laboratory conditions. This model allowed us to estimate the ambient illuminance during each child's tablet use based on the screen brightness data captured by the HSVC application. A controlled experiment was conducted in a professional lighting laboratory equipped with blackout conditions, where a response model between horizontal illuminance and screen brightness was constructed. The lighting environment was simulated using a warm-white light source with a correlated color temperature of 3000 K and a color rendering index (Ra) >80, to reflect typical nontask lighting conditions found in residential settings. The relative spectral power distribution of the light source is provided in Multimedia Appendix 1. This setup was designed to ensure both environmental realism and comfort.

The self-luminous device was operated under its default automatic brightness adjustment mode and fixed on a standard platform, with the screen center positioned at a height of 75 cm above the ground. A CL-500A spectroradiometer (KONICA MINOLTA) was used to record illuminance levels, with the sensor placed horizontally at the center of the desktop, approximately 30 cm in front of the screen. In accordance with the BS ISO/CIE 8995-1:2025 standard, horizontal illuminance was adopted as the primary photometric parameter due to its experimental stability and representativeness in evaluating screen-based visual environments. A total of 32 paired measurements of screen brightness and horizontal illuminance were collected by systematically adjusting ambient light intensity. A significant nonlinear relationship was observed between screen brightness and horizontal illuminance. As illustrated in Multimedia Appendix 2, the fitted quadratic regression model was as follows:

E=0.199x2-11.862x+177.930

where x denotes the screen brightness (unit: cd/m²) and E represents the corresponding horizontal illuminance (unit: lux). The model exhibited a high goodness of fit (R^2=0.988).

On the basis of the illuminance requirements specified in BS ISO/CIE 8995-1:2025 for *computer work only and young children* (300 lux) and *general classroom activities* (500 lux), the experimental data were classified into three ambient lighting categories: low illuminance (<300 lux), representing relaxed or recreational settings; moderate illuminance (300 - 500 lux), suitable for light visual tasks or display-based activities; and high illuminance (>500 lux), corresponding to functionally enhanced or brightly lit environments.

Covariates

A self-administered online questionnaire was distributed to parents to collect demographic information, including the child's age, sex, caregiver characteristics, only-child status, parental education levels, and household income. Additionally, the questionnaire collected relevant covariates based on existing literature, such as parental myopia status and children's daily behaviors (eg, duration of indoor and outdoor activities and sedentary time). Furthermore, the questionnaire assessed screen time across all media types, addressing the limitations of the HSVC in capturing data from television, personal computers, and other nontablet devices.

Statistical Analysis

Categorical variables were analyzed through frequency distributions (%), whereas continuous variables were presented according to their distribution. The normality of all continuous variables was assessed using the Shapiro-Wilk test. The Shapiro-Wilk test indicated that axial length and height were normally distributed (*P*>.05), whereas the other continuous variables deviated significantly from normality. Normally distributed data were presented as mean (SD), and nonnormally distributed data were presented as median (25th percentile and 75th percentile). Univariate analyses used independent *t* tests for binary comparisons (eg, sex differences) and 1-way ANOVA for multigroup comparisons (eg, age category). To evaluate monotonic relationships between variables, the Spearman rank correlation coefficient (ρ) was used. Spearman correlation coefficients quantified bivariate associations between axial length and continuous physiological or behavioral measures (eg, height, weight, and activity durations). Before model fitting, key regression assumptions were verified. The homoscedasticity assumption was evaluated using a Breusch-Pagan test, and the normality of residuals was formally assessed using the Shapiro-Wilk test; no substantial violations were detected (*P*>.05). Multivariable linear regression models adjusted for covariates (eg, age, sex, parental myopia, screen time, and activity levels) to assess independent effects of ambient illuminance and tablet brightness on axial length. Restricted cubic spline (RCS) models with 3 knots explored nonlinear dose-response relationships. Statistical significance was set at *P*<.05 (2 tailed), and analyses were performed using R software (version 4.3.2; R Foundation for Statistical Computing).

Ethical Considerations

Before the participants underwent screen exposure monitoring and questionnaire surveys, we informed their parents (legal guardians) of the entire process of participation in the surveys as well as the benefits and risks involved, and the legal guardians signed the informed consent. As a participation acknowledgment, children received art supply kits valued at US $15 per set. This study conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the Ethics Committee for Medical Research at the School of Public Health, Fudan University (IRB#2023-11-1088). When extracting data from the Huawei Research Platform and the online questionnaire library for analysis, all data were deidentified, and an anonymous study ID was used as an identifier for each participant.

Results

Participant Characteristics

Complete demographic and axial length data were available for all 199 participants. As presented in Table 1, 62.3% (n=124) of the participants were boys, and the majority (n=144, 72.4%) were aged between 4 years 9 months and 5 years 6 months. The participants' caregivers were predominantly their mothers (n=106, 53.3%), and more than half of the parents held a bachelor's degree (fathers: n=113, 56.8%; and mothers: n=113, 56.8%). The prevalence of maternal myopia was significantly higher (n=113, 56.8%) compared to paternal myopia (n=108, 54.3%), and maternal myopia was found to be significantly associated with a longer axial length (*P*=.01).

Table . Demographic characteristics and comparison of axial length between groups (N=199).

Variable	Values, n (%)	Axial length (mm), mean (SD)	*T* test/*F* test[a] (*df*)	*P* value[b]
Sex			6.184 (197)	<.001
Boy	124 (62.3)	22.68 (0.653)		
Girl	75 (37.7)	22.07 (0.717)		
Age			3.163 (2)	.04
4 y 9 mo to 5 y 6 mo	144 (72.4)	22.52 (0.747)		
4 y 3 mo to 4 y 9 mo	38 (19.0)	22.32 (0.682)		
3 y 3 mo to 4 y 3 mo	17 (8.5)	22.11 (0.692)		
Caregiver			1.254 (3)	.29
Mother only	106 (53.3)	22.52 (0.789)		
Father only	11 (5.5)	22.58 (0.749)		
Grandparents	78 (39.2)	22.33 (0.662)		
Other family members	4 (2.0)	22.67 (0.664)		
Education of father			1.571 (2)	.21
Secondary school or less	22 (11.0)	22.30 (0.755)		
Senior high school	64 (32.2)	22.36 (0.768)		
Bachelor’s	113 (56.8)	22.53 (0.715)		
Education of mother			1.065 (2)	.35
Secondary school or less	32 (16.1)	22.27 (0.742)		
Senior high school	54 (27.1)	22.48 (0.812)		
Bachelor’s	113 (56.8)	22.48 (0.700)		
Total annual household income (CNY)			–1.503 (197)	.13
<200,000 (US $27,800)	84 (42.2)	22.36 (0.719)		
≥200,000 (US $27,800)	115 (57.8)	22.52 (0.749)		
Only-child status			–0.389 (197)	.70
Yes	101 (50.8)	22.43 (0.714)		
No	98 (49.2)	22.47 (0.766)		
Current marital status of parents			–1.697 (197)	.09
Married	192 (96.5)	22.43 (0.734)		
Single/long-term separation	7 (3.5)	22.91 (0.767)		
Whether the father is nearsighted			–0.498 (197)	.62
No	91 (45.7)	22.42 (0.735)		
Yes	108 (54.3)	22.47 (0.744)		
Whether the mother is nearsighted			–2.470 (197)	.01
No	86 (43.2)	22.30 (0.750)		

Variable	Values, n (%)	Axial length (mm), mean (SD)	*T* test/*F* test[a] (*df*)	*P* value[b]
Yes	113 (56.8)	22.56 (0.713)		
Whether the child was ill in the past year			–1.167 (197)	.24
Yes	173 (86.9)	22.42 (0.740)		
No	26 (13.1)	22.61 (0.721)		

[a] *T* tests were used for binary comparisons, and the *F* statistic from a 1-way ANOVA was used for multigroup comparisons.

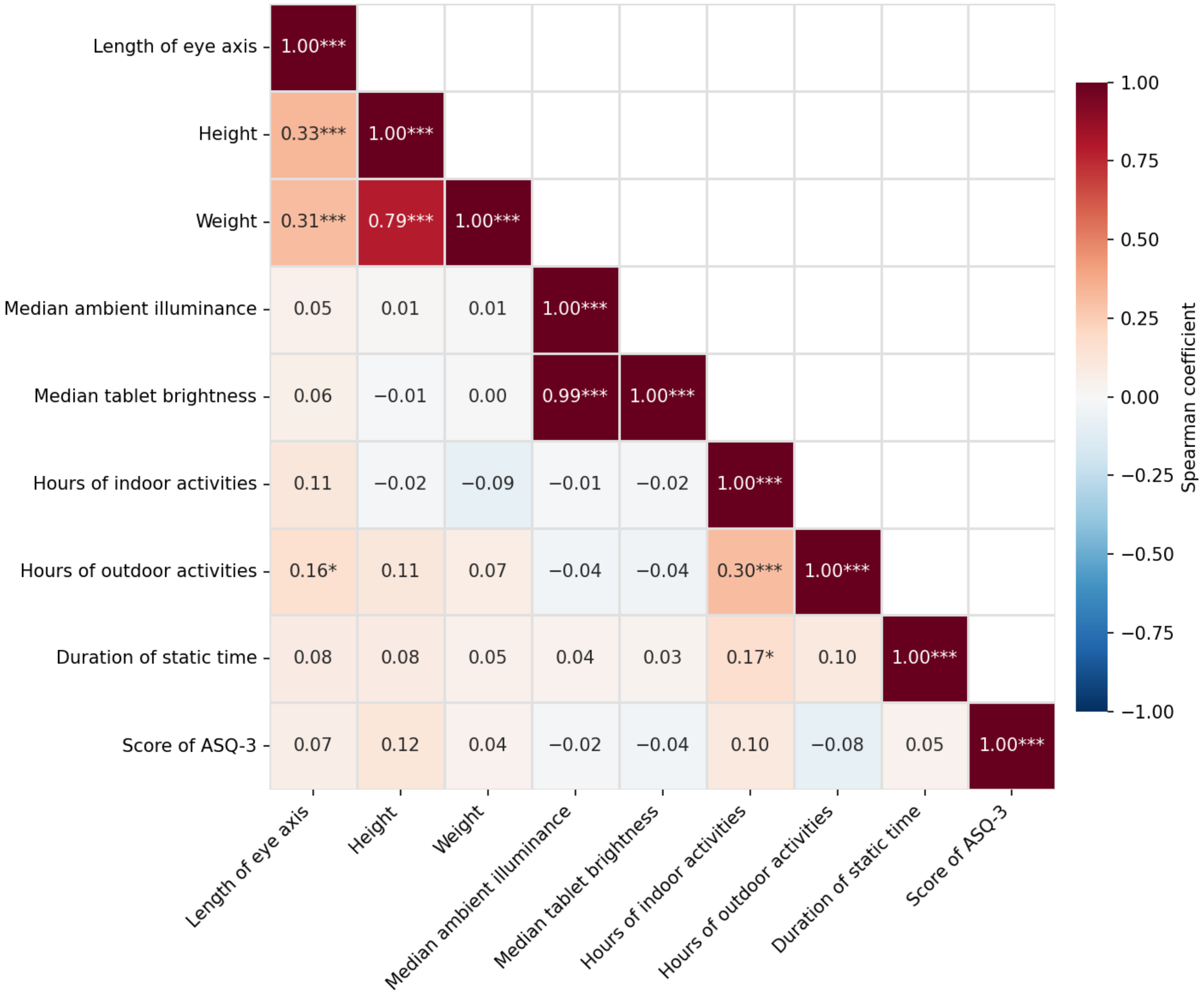
[b] *P* values were derived from independent *t* tests (for binary variables such as sex) or 1-way ANOVA (for multicategory variables such as age group). *P*<.05 was considered statistically significant.

Univariate analyses revealed significant differences in axial length based on sex, with boys having a greater mean axial length (22.68, SD 0.653 mm) compared to girls (22.07, SD 0.717 mm; *P*<.001). Additionally, significant differences in axial length were observed across different age groups (*P*=.04). However, no significant associations were found between axial length and variables, such as caregiver type, parental education level, household income, or only-child status (*P*>.05).

Correlation Analyses

The Spearman rank correlation analysis revealed several significant associations between certain demographic and behavioral factors and axial length. Specifically, height (ρ=0.33; *P*<.001), weight (ρ=0.31; *P*<.001), and outdoor activity duration (ρ=0.16; *P*=.03), all demonstrated statistically significant positive correlations with axial length (Figure 1). Notably, these relationships persisted even after adjusting for potential confounders, including parental myopia status and digital device exposure parameters.

Figure 1. Spearman rank correlation heatmap of key variables and axial length (N=199). Note: **P*<.05, ***P*<.01, ****P*<.001. ASQ-3: Ages and Stages Questionnaires, Third Edition.



Associations Between Ambient Illuminance or Tablet Brightness and Axial Length

As presented in Tables 2 and 3, multivariable linear regression demonstrated that higher median ambient illuminance during tablet use was inversely associated with axial length (β=–0.13, 95% CI –0.22 to –0.04; *P*=.006), after adjusting for sex, age, height, parental myopia, total screen time, and behavioral factors (Table 2). Specifically, the median ambient illuminance was 134.77 lux (P25=67.74 lux and P75=221.65 lux). Boy sex (*P*<.001) and greater height (*P*=.33) also significantly predicted axial length. However, there was no significant association between tablet brightness and axial length (Table 3). Although the median tablet brightness was 55 cd/m^2 (P25=47 cd/m^2 and P75=63.25 cd/m^2), it did not exhibit a linear association with axial length (Table 3).

Table . Associations between ambient illuminance and axial length (N=199)[a].

Variables	β[b] (95% CI)	*P* value
Height (cm)	0.16 (0.01 to 0.31)	.03
Weight (kg)	0.05 (–0.08 to 0.18)	.46
Sex		
Boy	Reference	—[c]
Girl	–0.59 (–0.79 to –0.40)	<.001
Age		
4 y 9 mo to 5 y 6 mo	Reference	—
4 y 3 mo to 4 y 9 mo	–0.01 (–0.28 to 0.25)	.92
3 y 3 mo to 4 y 3 mo	–0.12 (–0.51 to 0.26)	.52
Whether the mother is nearsighted		
No	Reference	—
Yes	0.12 (–0.07 to 0.31)	.20
Whether the father is nearsighted		
No	Reference	—
Yes	0.09 (–0.09 to 0.28)	.33
Total screen time used (ks)	0.03 (–0.06 to 0.12)	.56
Median ambient illuminance (lux)	–0.13 (–0.22 to –0.04)	.006
Hours of indoor activities (min)	0.05 (–0.05 to 0.14)	.31
Hours of outdoor activities (min)	0.05 (–0.04 to 0.14)	.26
Duration of static time (min)	–0.01 (–0.10 to 0.08)	.85
Score of ASQ-3[d] (pts)	0.07 (–0.02 to 0.16)	.13

[a] Results are derived from multivariable linear regression models. The dependent variable is axial length (mm).

[b] β represents the regression coefficient.

[c] Not applicable.

[d] ASQ-3: Ages & Stages Questionnaires, Third Edition.

Table . Associations between tablet brightness and axial length (N=199)[a].

Variables	β[b] (95% CI)	*P* value
Height (cm)	0.17 (0.02 to 0.32)	.02
Weight (kg)	0.04 (–0.09 to 0.17)	.51
Sex		
Boy	Reference	—
Girl	–0.58 (–0.78 to –0.39)	<.001
Age		
4 y 9 mo to 5 y 6 mo	Reference	—
4 y 3 mo to 4 y 9 mo	–0.01 (–0.28 to 0.26)	.94
3 y 3 mo to 4 y 3 mo	–0.10 (–0.49 to 0.29)	.60
Whether the mother is nearsighted		
No	Reference	—
Yes	0.12 (–0.07 to 0.31)	.22
Whether the father is nearsighted		
No	Reference	—
Yes	0.10 (–0.09 to 0.29)	.32
Total screen time used (ks)	0.03 (–0.06 to 0.12)	.51
Median tablet brightness (cd/m²)	–0.08 (–0.17 to 0.01)	.09
Hours of indoor activities (min)	0.05 (–0.05 to 0.14)	.35
Hours of outdoor activities (min)	0.06 (–0.04 to 0.15)	.24
Duration of static time (min)	–0.02 (–0.11 to 0.07)	.67
Score of ASQ-3[c] (pts)	0.07 (–0.03 to 0.16)	.17

[a] Results are derived from multivariable linear regression models. The dependent variable is axial length (mm).

[b] β represents the regression coefficient.

[c] ASQ-3: Ages and Stages Questionnaires, Third Edition.

Dose-Response Relationships Between Ambient Illuminance or Tablet Brightness and Axial Length

The dose-response relationships of tablet brightness and ambient illuminance with axial length were examined using RCS, as shown in Figures 2 and 3. Multivariable linear regression models incorporating RCS were fitted with 3 knots placed at the 10th, 50th, and 90th percentiles of the exposure distribution [17]. All models were adjusted for sociodemographic factors (eg, boy or girl, age, height, weight, and parental myopia), behavioral factors (indoor and outdoor activity duration, and sedentary time), and developmental characteristics (score from Ages and Stages Questionnaires, Third Edition). A significant nonlinear association was identified between median tablet brightness and axial length ($P_{nonlinearity}$=.004). The coefficient β represents the marginal change in axial length (in mm) per unit increase in exposure. The fitted curve indicated a negative slope at brightness levels below 27 cd/m², suggesting a reduction in the rate of axial elongation. This was followed by a transitional phase (27 ‐ 56 cd/m²) where the curve became positive and increased, peaking at approximately 56 cd/m². Beyond this point, the fitted curve gradually decreased. In contrast, no significant nonlinear relationship was found between ambient illuminance and axial length ($P_{nonlinearity}$=.17), with the curve remaining approximately flat across the measured range (Figure 3).

Figure 2. Dose-response relationship between tablet brightness and axial length (N=199). The solid line represents the estimated adjusted regression coefficient β, and the blue area indicates the 95% CI. Knots were placed at the 10th, 50th, and 90th percentiles of the median tablet brightness.

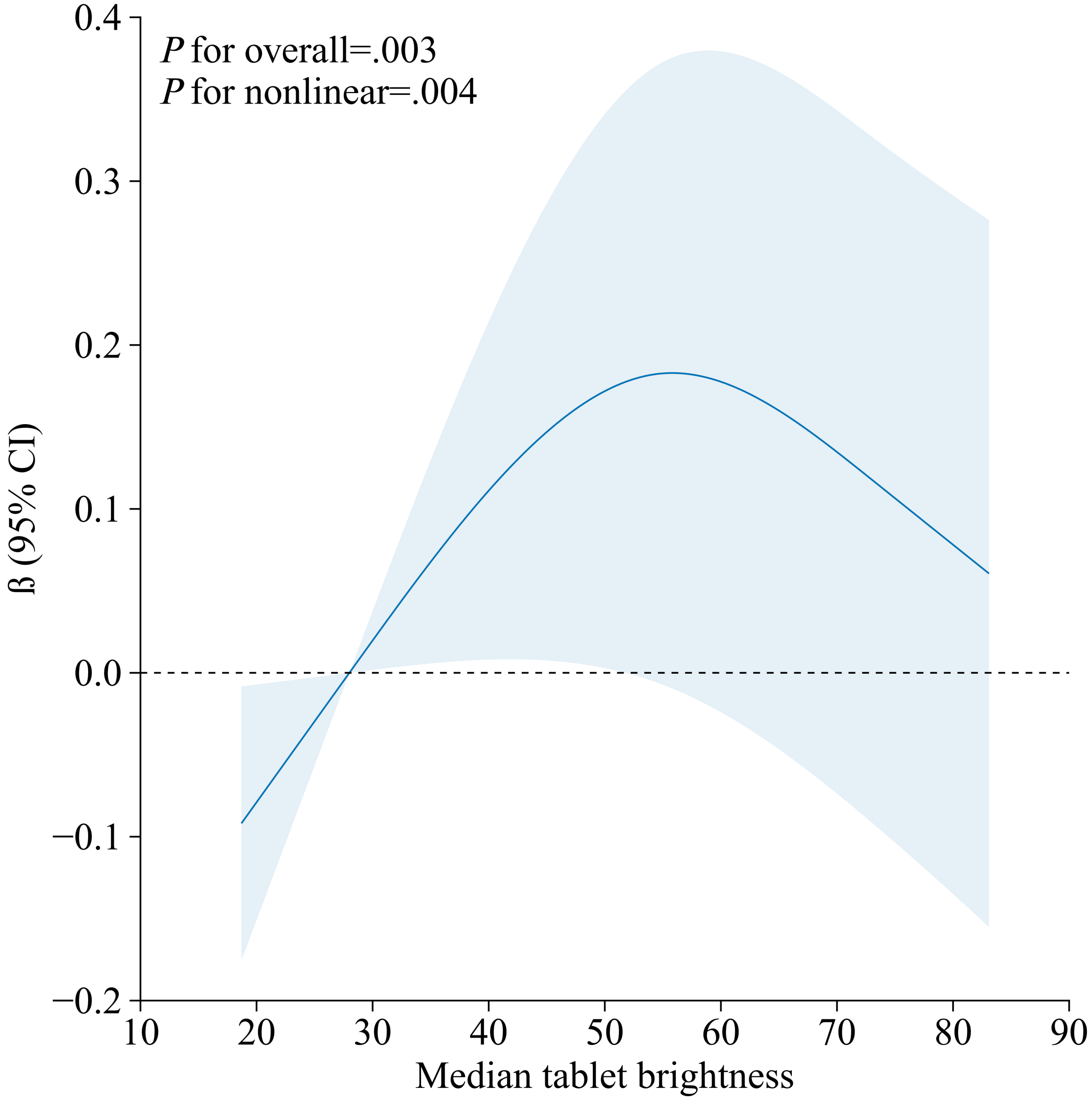
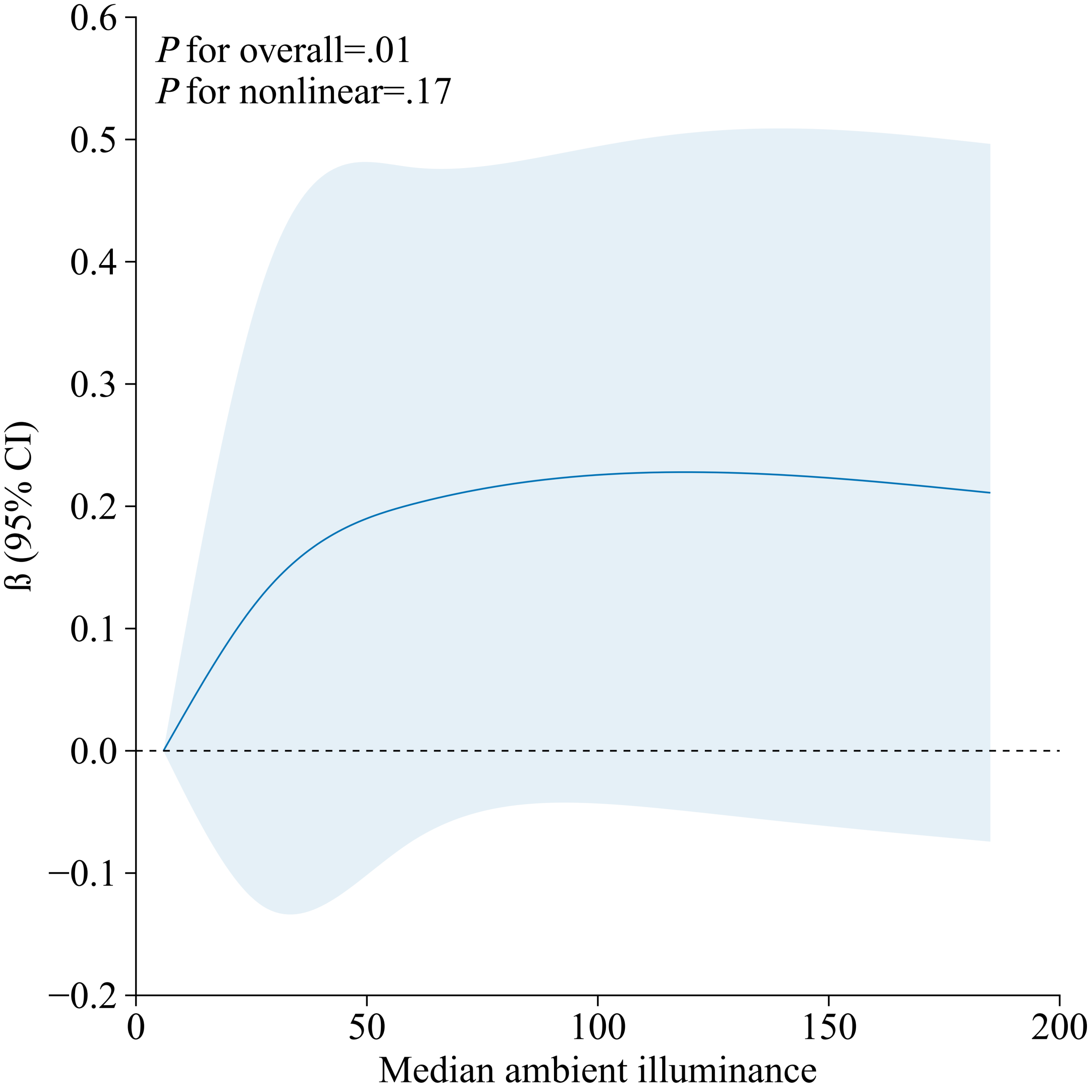


Figure 3. Dose-response relationship between ambient illuminance and axial length (N=199). The solid line represents the estimated adjusted regression coefficient β, and the blue area indicates the 95% CI. Knots were placed at the 10th, 50th, and 90th percentiles of the median ambient illuminance.



Discussion

Principal Findings

On the basis of continuous 7-day real-time monitoring data collected via the HSVC application from 199 preschool children aged 3 to 6 years in 2 kindergartens in Shanghai, this study investigated the associations among screen luminance, ambient illuminance during device use, and axial length in early childhood, addressing a critical gap in the existing literature regarding the impact of dynamic screen optical parameters on visual development in children. Our findings reveal a significant inverse association between higher ambient illuminance during tablet use and shorter axial length, as well as a nonlinear dose-response relationship between tablet luminance and axial length. These results elucidate the link between digital device optical exposure and axial elongation, providing robust empirical evidence to inform public health recommendations on screen use practices and ambient lighting standards in preschool children.

Relationship Between Ambient Illuminance and Axial Length

Our study identified a significant inverse association between ambient illuminance during tablet use and axial length in preschool children, aligning with findings from animal models [14]. Experimental studies in chicks and rhesus monkeys have established that light intensity modulates myopic progression [18-21]. Multiple investigations further demonstrate that increased light exposure may exert its effect by activating retinal dopaminergic activity—primarily through D2 receptor pathways—thereby regulating retinal gene expression and suppressing pro-axial elongation signaling [22-25]. The retinal dopamine system is thus considered a potent inhibitor of ocular growth [26].

The present results align with several domestic studies investigating the role of light environment in myopia prevention. For instance, a study on elevated classroom lighting found that increasing illuminance significantly slowed axial elongation and reduced the incidence of myopia among school-aged children [27]. Similarly, a cluster randomized trial in Shanghai using objective monitoring showed that increased outdoor time and sustained high-intensity light exposure delayed axial elongation [28]. Together, these studies support a protective role of light exposure against axial growth. Our study extends this evidence from "macroenvironmental" settings to "microenvironments during screen use," refining the understanding of light exposure effects.

Complementing this mechanistic context, it is important to recognize that not all studies of indoor optical environments have detected an effect on refractive development. For instance, a school-based comparison of high-daylight versus low-daylight classrooms reported no overall difference in 6-month changes in refractive error or axial length, with effects limited only to a baseline subgroup characterized by short axial length [29]. Similarly, an independent population-based analysis found no association between exposure to ambient nighttime lighting during infancy and later myopia [30]. These inconsistent outcomes may be attributable to methodological and contextual differences. The former study used classroom-level daylight factors—a distal and nonspecific metric—which may not accurately capture retinal light exposure during near-device use. Furthermore, the study was conducted in a specific ethnic and educational setting (Korean schoolchildren), where genetic predispositions and academic pressures may modulate the effects of light exposure. The latter study examined nighttime lighting exposure in infancy. This differs fundamentally from daytime ambient light during visual tasks in its spectral composition, exposure timing, and biological mechanisms. Furthermore, the cohort primarily comprised White participants, who have a distinct genetic background from Asian populations. These methodological and demographic differences may collectively account for the heterogeneity observed regarding the association between light exposure and myopia.

It is noteworthy that typical indoor illuminance experienced by humans rarely exceeds 1000 lux and is usually much lower, often within the range of 100 to 500 lux [31]. According to standards set by the International Organization for Standardization and the International Commission on Illumination, as well as China's national guidelines (GB 7793 - 2010) on classroom lighting, the recommended threshold for children's indoor ambient illuminance is 300 lux [32,33]. In our findings, the median ambient illuminance during tablet use for 174 children was below 300 lux, with an overall median of 134.77 (IQR 67.74-221.65) lux, indicating that the ambient lighting conditions when participants used electronic devices were generally lower than internationally recommended standards. Nevertheless, we still observed a robust association between ambient illuminance and axial length, underscoring the critical role of environmental lighting in ocular development. Suboptimal lighting may fail to provide adequate visual comfort and could contribute to visual fatigue or other adverse effects on visual development. The widespread prevalence of insufficient ambient lighting during device use is concerning, as it represents a potentially modifiable environmental risk factor for myopia in children. Given the increasing reliance on digital devices in early childhood, our findings highlight the urgent need for evidence-based guidelines on ambient lighting for children to mitigate potential risks to visual health and development.

However, our results diverge from some previous reports, suggesting that ambient illuminance has no significant effect on visual fatigue [34]. This discrepancy suggests that the mechanisms underlying short-term visual discomfort and long-term regulation of axial growth may differ and warrant separate investigation. While short-term visual fatigue is likely associated with transient visual stimuli and ocular muscle strain, long-term axial elongation is influenced by a complex interplay of factors, including light-mediated regulation of ocular signaling pathways. From the perspective of long-term axial development, our study provides novel evidence for the impact of ambient illuminance on children's visual health.

Relationship Between Tablet Brightness and Axial Length

Using RCS with knots at the 10th, 50th, and 90th percentiles of the exposure distribution, we identified a significant nonlinear association between median tablet brightness and axial length after adjustment for sociodemographic, behavioral, and developmental covariates ($P_{\text{nonlinearity}}$=.004). At lower levels of tablet brightness (with a median tablet brightness <27 cd/m²), increases in brightness are accompanied by a deceleration in axial length growth (β<0), suggesting that very dim screens may accelerate eye growth, whereas moderate brightness increases could help mitigate excessively rapid axial elongation. When tablet brightness exceeds a certain threshold (median tablet brightness >27 cd/m²), further increases in brightness result in an acceleration of axial elongation, with excessively high screen brightness stimulating abnormal axial growth. When tablet brightness reaches a critical threshold (median tablet brightness=56 cd/m²), the rate of axial elongation peaks. Excessive brightness can cause glare or strong light stimuli, disrupting normal physiological eye development and accelerating axial elongation, which may worsen myopia progression [13]. In this study, the median tablet brightness was 56 cd/m² (P25=47 cd/m² and P75=63.25 cd/m²), falling within the critical range that may influence ocular growth.

These findings align with the "parabolic or inverted U-shaped dose-response curve" described in previous ergonomic studies regarding screen brightness, visual performance, and discomfort symptoms [15]. At very low brightness, children may reduce viewing distance to maintain visibility, increasing accommodative demand and potentially stimulating axial elongation. Moderately increased brightness may improve legibility and allow for more comfortable viewing distances, corresponding to the observed negative marginal effect below 27 cd/m² [35,36]. Within the midrange, heightened brightness during sustained near-work may provide stronger visual signals associated with axial elongation, resulting in a positive marginal effect that peaks near 56 cd/m² [35,37]. At even higher brightness, discomfort and incipient glare may prompt self-regulatory behaviors—such as microbreaks, blinking, or adjustment of posture or gaze—that could attenuate the marginal effect, consistent with the decline in the fitted curve at high brightness levels [37-39]. This nonmonotonic pattern aligns with ergonomic literature in which display brightness often exhibits nonlinear relationships with both performance and discomfort [15,40]. To our knowledge, this is the first study to quantitatively characterize a nonlinear dose-response relationship between objectively measured tablet brightness and axial length in children, addressing a significant evidence gap in this field. This finding is also consistent with the “light dose-axial plasticity” effect reported in domestic studies on optical interventions. For instance, a randomized clinical trial in 10 primary schools in Shanghai demonstrated that repeated low-level red light therapy reduced the annual axial elongation rate in children from 0.47 mm to 0.30 mm over 12 months, alongside a significant decrease in new-onset myopia [41]. Although repeated low-level red light represents a therapeutic intervention fundamentally distinct from daily screen-based illuminance, its outcomes clearly indicate that the mode and dosage of light stimulation can structurally influence axial elongation. This lends biological plausibility to the nonlinear association between screen brightness and axial length observed under naturalistic conditions in our study and further supports the notion that optical parameters play a modifiable and nonlinear role in axial eye growth.

More importantly, the study identifies a relatively “safe” brightness range for children’s use of mobile devices, within which axial elongation remains stable and the impact on myopia progression is minimal. To translate these key findings into practical application, device manufacturers should implement child-specific brightness-limiting features that cap maximum screen luminance below 56 cd/m², incorporating ambient-adaptive safeguards to prevent exceedance of this threshold and thus reduce the risk of excessive axial elongation. Caregivers should be advised to maintain screen brightness within the empirically derived safe range identified in this study and to encourage regular, timed breaks during prolonged use to minimize potential adverse effects on visual development. These results provide robust scientific evidence to support public health guidelines for optimal screen brightness settings in both educational and household environments.

Limitations

Several limitations of this study should be acknowledged. First, the cross-sectional nature of this study precludes any causal inferences. Future prospective studies are needed to determine whether reduced ambient illuminance directly accelerates axial elongation and to clarify causal relationships among relevant factors. Second, our sample was limited to children in Shanghai, potentially restricting the generalizability of our findings to populations with different screen usage habits. Future research should extend to larger and more diverse cohorts across various regions and populations to enhance external validity. Additionally, although we attempted to control for covariates, such as the duration of outdoor activities and unmeasured confounders—including genetic predisposition and dietary factors—may partially account for the observed associations. Future research should address these factors to comprehensively elucidate the determinants affecting axial eye development in children. Finally, as cycloplegic refraction was not performed in this study, the accuracy of refractive error measurements was limited. Therefore, axial length was selected as the primary outcome. In future research, we will refine the methodology by incorporating cycloplegic refraction to enable a more comprehensive evaluation of the effects of ambient illuminance on children’s visual health.

Conclusions

This monitoring study demonstrates a significant association between both ambient illuminance and tablet screen brightness with ocular axial length in preschool children, with a nonlinear dose-response relationship observed for screen brightness. These findings underscore the importance of optimizing both environmental lighting and device settings to safeguard early childhood visual development and inform evidence-based guidelines for digital device use in preschool children.

Acknowledgments

The authors would like to extend our sincere appreciation to Huawei for developing the intelligent monitoring software used to assess children’s screen exposure in this study. The authors are also deeply grateful to the kindergarten teachers from Pudong New District and Xuhui District for their invaluable assistance, as well as to all the participants and their families for their contributions to this research.

Funding

This work was supported by the National Natural Science Foundation of China (grant 82304260); Key Discipline and Project of High-Quality Development of Public Health of School of Public Health, Fudan University–Jiading District Health Commission (grant GWGZLXK-2023 04); and Shanghai Municipal Health Commission (grant GWV-10.2-YQ23). The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

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

Authors' Contributions

Conceptualization: YJ (lead), HC (equal), HJ (supporting), SL (supporting), YZ (supporting)
Formal analysis: HC (lead), YZ (equal), MZ (supporting)
Funding acquisition: YJ (lead), HC (equal), HJ (supporting)
Investigation: YZ (lead), HC (equal), YS (supporting), CL (supporting), MZ (supporting)
Methodology: YJ (lead), HC (equal), HJ (supporting), SL (supporting), YZ (supporting)
Project administration: YJ (lead), MZ (equal), HJ (supporting)
Resources: YJ (lead), MZ (equal), HJ (supporting)
Supervision: YJ (lead), HJ (equal), SL (supporting), HC (supporting)
Writing – original draft: YZ (lead), MZ (supporting)
Writing – review & editing: YJ (lead), HC (equal), HJ (supporting), SL (supporting), YZ (supporting), MZ (supporting), YS (supporting), CL (supporting)
HJ is the senior author for this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Relative spectral power distribution of the light source (correlated color temperature=3000 K).
[DOCX File, 84 KB - humanfactors_v13i1e79266_app1.docx]

Multimedia Appendix 2
Nonlinear response relationship between screen luminance and horizontal illuminance of the self-luminous device.
[DOCX File, 93 KB - humanfactors_v13i1e79266_app2.docx]

References

1. Dolgin E. The myopia boom. Nature 2015 Mar 19;519(7543):276-278. [doi: 10.1038/519276a] [Medline: 25788077]
2. Chua SYL, Sabanayagam C, Cheung YB, et al. Age of onset of myopia predicts risk of high myopia in later childhood in myopic Singapore children. Ophthalmic Physiologic Optic 2016 Jul;36(4):388-394. [doi: 10.1111/opo.12305]
3. Pärssinen O, Kauppinen M. Risk factors for high myopia: a 22-year follow-up study from childhood to adulthood. Acta Ophthalmol 2019 Aug;97(5):510-518. [doi: 10.1111/aos.13964] [Medline: 30460746]
4. Haarman AEG, Enthoven CA, Tideman JWL, Tedja MS, Verhoeven VJM, Klaver CCW. The complications of myopia: a review and meta-analysis. Invest Ophthalmol Vis Sci 2020 Apr 9;61(4):49. [doi: 10.1167/iovs.61.4.49] [Medline: 32347918]
5. Tremblay MS, Barnes JD, González SA, et al. Global matrix 2.0: report card grades on the physical activity of children and youth comparing 38 countries. J Phys Act Health 2016 Nov;13(11 Suppl 2):S343-S366. [doi: 10.1123/jpah.2016-0594] [Medline: 27848745]
6. Rideout V, Robb MB. The common sense census: media use by kids age zero to eight, 2020. : Common Sense Media; 2020 URL: https://www.commonsensemedia.org/sites/default/files/research/report/2020_zero_to_eight_census_final_web.pdf [accessed 2025-12-10]
7. Tooth LR, Moss KM, Mishra GD. Adherence to screen time guidelines among families in Australia with children of different ages. JAMA Pediatr 2022 May 1;176(5):517-519. [doi: 10.1001/jamapediatrics.2021.6382] [Medline: 35157014]
8. Foreman J, Salim AT, Praveen A, et al. Association between digital smart device use and myopia: a systematic review and meta-analysis. Lancet Digit Health 2021 Dec;3(12):e806-e818. [doi: 10.1016/S2589-7500(21)00135-7] [Medline: 34625399]
9. Terasaki H, Yamashita T, Yoshihara N, Kii Y, Sakamoto T. Association of lifestyle and body structure to ocular axial length in Japanese elementary school children. BMC Ophthalmol 2017 Jul 12;17(1):123. [doi: 10.1186/s12886-017-0519-y] [Medline: 28697750]
10. Enthoven CA, Polling JR, Verzijden T, et al. Smartphone use associated with refractive error in teenagers. Ophthalmology 2021 Dec;128(12):1681-1688. [doi: 10.1016/j.ophtha.2021.06.016] [Medline: 34245754]
11. Benedetto S, Carbone A, Drai-Zerbib V, Pedrotti M, Baccino T. Effects of luminance and illuminance on visual fatigue and arousal during digital reading. Comput Human Behav 2014 Dec;41:112-119. [doi: 10.1016/j.chb.2014.09.023]
12. Rosenfield M. Computer vision syndrome: a review of ocular causes and potential treatments. Ophthalmic Physiol Opt 2011 Sep;31(5):502-515. [doi: 10.1111/j.1475-1313.2011.00834.x] [Medline: 21480937]
13. Yu H, Akita T, Koga T, Sano N. Effect of character contrast ratio of tablet PC and ambient device luminance ratio on readability in low ambient illuminance. Displays 2018 Apr;52:46-54. [doi: 10.1016/j.displa.2018.03.002]
14. Norton TT. What do animal studies tell us about the mechanism of myopia—protection by light? Optom Vis Sci 2016 Sep;93(9):1049-1051. [doi: 10.1097/OPX.0000000000000917] [Medline: 27362614]

15. Lin C, Ji Z, Lin Y. Optimum display luminance and contrast polarity of desktop head-up display under office lighting level based on visual ergonomic study. Ergonomics 2024 Nov;67(11):1491-1503. [doi: 10.1080/00140139.2024.2339439] [Medline: 38629123]
16. Chen H, Sun Y, Luo S, et al. Association of screen content with early development among preschoolers in Shanghai: 7-day monitoring study with auto intelligent technology. J Med Internet Res 2025 Mar 5;27:e65343. [doi: 10.2196/65343] [Medline: 40053802]
17. Austin PC, Fang J, Lee DS. Using fractional polynomials and restricted cubic splines to model non-proportional hazards or time-varying covariate effects in the cox regression model. Stat Med 2022 Feb 10;41(3):612-624. [doi: 10.1002/sim.9259] [Medline: 34806210]
18. Karouta C, Ashby RS. Correlation between light levels and the development of deprivation myopia. Invest Ophthalmol Vis Sci 2014 Dec 9;56(1):299-309. [doi: 10.1167/iovs.14-15499] [Medline: 25491298]
19. Smith EL III, Hung LF, Huang J. Protective effects of high ambient lighting on the development of form-deprivation myopia in rhesus monkeys. Invest Ophthalmol Vis Sci 2012 Jan 25;53(1):421-428. [doi: 10.1167/iovs.11-8652] [Medline: 22169102]
20. Lan W, Feldkaemper M, Schaeffel F. Intermittent episodes of bright light suppress myopia in the chicken more than continuous bright light. PLoS One 2014;9(10):e110906. [doi: 10.1371/journal.pone.0110906] [Medline: 25360635]
21. Rucker F, Britton S, Taylor C. Color and Temporal Frequency Sensitive Eye Growth in Chick. Invest Ophthalmol Vis Sci 2018 Dec 3;59(15):6003-6013. [doi: 10.1167/iovs.18-25322] [Medline: 30572345]
22. Stone RA, Lin T, Laties AM, Iuvone PM. Retinal dopamine and form-deprivation myopia. Proc Natl Acad Sci U S A 1989 Jan;86(2):704-706. [doi: 10.1073/pnas.86.2.704] [Medline: 2911600]
23. Iuvone PM, Tigges M, Fernandes A, Tigges J. Dopamine synthesis and metabolism in rhesus monkey retina: development, aging, and the effects of monocular visual deprivation. Vis Neurosci 1989;2(5):465-471. [doi: 10.1017/s0952523800012360] [Medline: 2577263]
24. McCarthy CS, Megaw P, Devadas M, Morgan IG. Dopaminergic agents affect the ability of brief periods of normal vision to prevent form-deprivation myopia. Exp Eye Res 2007 Jan;84(1):100-107. [doi: 10.1016/j.exer.2006.09.018] [Medline: 17094962]
25. Feldkaemper M, Schaeffel F. An updated view on the role of dopamine in myopia. Exp Eye Res 2013 Sep;114:106-119. [doi: 10.1016/j.exer.2013.02.007] [Medline: 23434455]
26. Iuvone PM, Tigges M, Stone RA, Lambert S, Laties AM. Effects of apomorphine, a dopamine receptor agonist, on ocular refraction and axial elongation in a primate model of myopia. Invest Ophthalmol Vis Sci 1991 Apr;32(5):1674-1677. [Medline: 2016144]
27. Hua WJ, Jin JX, Wu XY, et al. Elevated light levels in schools have a protective effect on myopia. Ophthalmic Physiol Opt 2015 May;35(3):252-262. [doi: 10.1111/opo.12207] [Medline: 25913873]
28. He X, Sankaridurg P, Wang J, et al. Time outdoors in reducing myopia: a school-based cluster randomized trial with objective monitoring of outdoor time and light intensity. Ophthalmology 2022 Nov;129(11):1245-1254. [doi: 10.1016/j.ophtha.2022.06.024] [Medline: 35779695]
29. Suh YW, Ha SG, Kim SH. Effect of classroom illuminance on the development and progression of myopia in school children. Korean J Ophthalmol 2022 Jun;36(3):194-201. [doi: 10.3341/kjo.2021.0170] [Medline: 35067020]
30. Gwiazda J, Ong E, Held R, Thorn F. Myopia and ambient night-time lighting. Nature 2000 Mar 9;404(6774):144-144. [doi: 10.1038/35004663] [Medline: 10724158]
31. Norton TT, Siegwart JT. Light levels, refractive development, and myopia--a speculative review. Exp Eye Res 2013 Sep;114:48-57. [doi: 10.1016/j.exer.2013.05.004] [Medline: 23680160]
32. BS ISO/CIE 8995-1:2025 light and lighting — lighting of work places — part 1: indoor. : International Organization for Standardization & CIE; 2025 URL: https://www.iso.org/standard/76342.html [accessed 2025-12-10]
33. GB/t 36876-2018. hygienic requirements of lighting design and setting in primary and middle schools. : State Administration for Market Regulation, Standardization Administration of China; 2018 URL: https://www.chinesestandard.us/products/gbt36876-2018 [accessed 2025-12-10]
34. Chen JG, Wu SK, Chiu HP, Tu CN, Liu CH. Evaluation of three tablet computers at two levels of ambient illumination. Int J Hum Comput Interact 2016 May 3;32(5):394-401. [doi: 10.1080/10447318.2016.1151614]
35. Zhou Y, Shi H, Chen QW, Ru T, Zhou G. Investigation of the optimum display luminance of an LCD screen under different ambient illuminances in the evening. Appl Sci 2021;11(9):4108. [doi: 10.3390/app11094108]
36. Shieh KK, Lee DS. Preferred viewing distance and screen angle of electronic paper displays. Appl Ergon 2007 Sep;38(5):601-608. [doi: 10.1016/j.apergo.2006.06.008] [Medline: 17049333]
37. Sheedy JE, Smith R, Hayes J. Visual effects of the luminance surrounding a computer display. Ergonomics 2005 Jul 15;48(9):1114-1128. [doi: 10.1080/00140130500208414] [Medline: 16251151]
38. Yin Z, Liu B, Hao D, Yang L, Feng Y. Evaluation of VDT-induced visual fatigue by automatic detection of blink features. Sensors (Basel) 2022 Jan 25;22(3):916. [doi: 10.3390/s22030916] [Medline: 35161662]
39. Miyake-Kashima M, Dogru M, Nojima T, Murase M, Matsumoto Y, Tsubota K. The effect of antireflection film use on blink rate and asthenopic symptoms during visual display terminal work. Cornea 2005 Jul;24(5):567-570. [doi: 10.1097/01.ico.0000151564.24989.38] [Medline: 15968162]

40. Muhamad N, Amali NAN. Digital display preference of electronic gadgets for visual comfort: a systematic review. Iran J Public Health 2023 Aug;52(8):1565-1577. [doi: 10.18502/ijph.v52i8.13396] [Medline: 37744537]
41. He X, Wang J, Zhu Z, et al. Effect of repeated low-level red light on myopia prevention among children in China with premyopia: a randomized clinical trial. JAMA Netw Open 2023 Apr 3;6(4):e239612. [doi: 10.1001/jamanetworkopen.2023.9612] [Medline: 37099298]

Abbreviations

HSVC: Healthy Screen Viewing for Children
RCS: restricted cubic spline

Edited by A Kushniruk; submitted 19.Jun.2025; peer-reviewed by R Doron, W Yang; revised version received 18.Nov.2025; accepted 01.Dec.2025; published 28.Jan.2026.

Please cite as:
Zhu Y, Chen H, Lin S, Jiang H, Zhang M, Sun Y, Li C, Jia Y
Objective Monitoring of Tablet Use–Related Optical Exposure and Its Association With Axial Length in Preschool Children: Cross-Sectional Intelligent Monitoring Study
JMIR Hum Factors 2026;13:e79266
URL: https://humanfactors.jmir.org/2026/1/e79266
doi:10.2196/79266

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org