Contents

Original Papers

Development of eHOME, a Mobile Instrument for Reporting, Monitoring, and Consulting Drug-Related Problems in Home Care: Human-Centered Design Study (e10)
Nienke Dijkstra, Carolina Sino, Eibert Heerdink, Marieke Schuurmans. ................................................................. 3

Reducing Misses and Near Misses Related to Multitasking on the Electronic Health Record: Observational Study and Qualitative Analysis (e4)
Neda Ratanawongsa, George Matta, Fuad Bohsali, Margaret Chisolm. ................................................................. 25

Second Version of Google Glass as a Wearable Socio-Affective Aid: Positive School Desirability, High Usability, and Theoretical Framework in a Sample of Children with Autism (e1)
Ned Sahin, Neha Keshav, Joseph Salisbury, Arshya Vahabzadeh. ................................................................. 31

How Can Social Media Lead to Co-Production (Co-Delivery) of New Services for the Elderly Population? A Qualitative Study (e5)
Hadi Daneshvar, Stuart Anderson, Robin Williams, Hajar Mozaffar. ................................................................. 43

VA FitHeart, a Mobile App for Cardiac Rehabilitation: Usability Study (e3)
Alexis Beatty, Sara Magnusson, John Fortney, George Sayre, Mary Whooley. ................................................................. 57

How Health Care Professionals Evaluate a Digital Intervention to Improve Medication Adherence: Qualitative Exploratory Study (e7)
Karen Thomson, Corline Brouwers, Olga Damman, Martine de Bruijne, Danielle Timmermans, Marijke Melles. ................................................................. 68

Evaluating the Usability and Usefulness of a Mobile App for Atrial Fibrillation Using Qualitative Methods: Exploratory Pilot Study (e13)
Jaclyn Hirschey, Sunetra Bane, Moussa Mansour, Jodi Sperber, Stephen Agboola, Joseph Kvedar, Kamal Jethwani. ................................................................. 82

The Perceived Ease of Use and Usefulness of Loop: Evaluation and Content Analysis of a Web-Based Clinical Collaboration System (e2)
Allison Kurahashi, Jennifer Stinson, Margaret van Wyk, Stephanie Luca, Trevor Jamieson, Peter Weinstein, Joseph Cafazzo, Bhadra Lokuge, Eyal Cohen, Adam Rapoport, Amna Husain. ................................................................. 93

Co-Designing a Collaborative Chronic Care Network (C3N) for Inflammatory Bowel Disease: Development of Methods (e8)
Michael Seid, George Dellal, Laura Peterson, Lloyd Provost, Peter Gloo, David Fore, Peter Margolis. ................................................................. 111
Supporting Accurate Interpretation of Self-Administered Medical Test Results for Mobile Health: Assessment of Design, Demographics, and Health Condition (e9)
Jess Hohenstein, Eric Baumer, Lindsay Reynolds, Elizabeth Murnane, Dakota O'Dell, Seoho Lee, Shion Guha, Yu Qi, Erin Rieger, Geri Gay.

General Practitioners’ Attitudes Toward a Web-Based Mental Health Service for Adolescents: Implications for Service Design and Delivery (e12)
Mirjana Subotic-Kerry, Catherine King, Kathleen O'Moore, Melinda Achilles, Bridianne O'Dea.

Three-Factor Structure of the eHealth Literacy Scale Among Magnetic Resonance Imaging and Computed Tomography Outpatients: A Confirmatory Factor Analysis (e6)
Lisa Hyde, Allison Boyes, Tiffany-Jane Evans, Lisa Mackenzie, Rob Sanson-Fisher.

Review

Human Factors and Data Logging Processes With the Use of Advanced Technology for Adults With Type 1 Diabetes: Systematic Integrative Review (e11)
Marion Waite, Clare Martin, Rachel Franklin, David Duce, Rachel Harrison.
Development of eHOME, a Mobile Instrument for Reporting, Monitoring, and Consulting Drug-Related Problems in Home Care: Human-Centered Design Study

Abstract

Background: Home care patients often use many medications and are prone to drug-related problems (DRPs). For the management of problems related to drug use, home care could add to the multidisciplinary expertise of general practitioners (GPs) and pharmacists. The home care observation of medication-related problems by home care employees (HOME)-instrument is paper-based and assists home care workers in reporting potential DRPs. To facilitate the multiprofessional consultation, a digital report of DRPs from the HOME-instrument and digital monitoring and consulting of DRPs between home care and general practices and pharmacies is desired.

Objective: The objective of this study was to develop an electronic HOME system (eHOME), a mobile version of the HOME-instrument that includes a monitoring and a consulting system for primary care.

Methods: The development phase of the Medical Research Council (MRC) framework was followed in which an iterative human-centered design (HCD) approach was applied. The approach involved a Delphi round for the context of use and user requirements analysis of the digital HOME-instrument and the monitoring and consulting system followed by 2 series of pilots for testing the usability and redesign.

Results: By using an iterative design approach and by involving home care workers, GPs, and pharmacists throughout the process as informants, design partners, and testers, important aspects that were crucial for system realization and user acceptance were revealed. Through the report webpage interface, which includes the adjusted content of the HOME-instrument and added home care practice–based problems, home care workers can digitally report observed DRPs. Furthermore, it was found that the monitoring and consulting webpage interfaces enable digital consultation between home care and general practices and pharmacies. The webpages were considered convenient, clear, easy, and usable.

Conclusions: By employing an HCD approach, the eHOME-instrument was found to be an easy-to-use system. The systematic approach promises a valuable contribution for the future development of digital mobile systems of paper-based tools.

KEYWORDS
primary care; home care; eHealth; mHealth

doi:10.2196/humanfactors.8319
**Introduction**

**Background**
Pharmacotherapy is one of the most common interventions used in health care. Its use has considerably grown because of the aging population and the increased prevalence of chronic diseases [1]. Although medication may contribute to cure, slow progression, or reduce the symptoms of diseases, it is also associated with drug-related problems (DRPs). According to the Pharmaceutical Care Network Europe (PCNE), a DRP is defined as an event or circumstance involving drug therapy that actually or potentially interferes with desired health outcomes [2]. DRPs may negatively affect a person’s perceived quality of life, and it may increase morbidity, mortality, health care costs, and the risk of hospital (re)admissions [3-6]. Older people are more prone to DRPs because of the higher prevalence of drug use and age-related pathophysiologic changes in pharmacokinetics and pharmacodynamics [3,7,8]. To prevent or limit complications of DRPs in older people, a multidisciplinary approach in primary care through home care, general practices, and pharmacies is desirable [9,10]. Home care workers have insight into the home environment of the patients and can offer an important contribution to the recognition of problems related to drug use. Several tools have been developed for home care workers for the recognition of signs and symptoms of DRPs in home care patients [11-15]. One of these instruments is the validated home care observation of medication-related problems by home care employees (HOME)-instrument [15], an observation list with 28 signs and symptoms of potential DRPs categorized in 3 categories (process, pill, and patient). An observational study of the HOME-instrument by our research group [15] showed that almost half of all the observed signs and symptoms were assessed as potentially drug related. The challenges of the HOME-instrument observation forms include issues such as problems to transfer or store forms into an electronic patient file and difficulty in monitoring and comparing the progression of DRPs over time. Furthermore, it does not offer consultation with other primary care disciplines, such as general practitioners (GPs) and pharmacists. Digital health care technologies (eHealth) combining information and communication technologies may eliminate these challenges. GPs and pharmacists expressed the demand for collaboration in recognizing and managing of DRPs by using the HOME-instrument as a mobile system. Collaboration between different health care professionals can be used to develop a usable electronic HOME system (eHOME) that combines the report of DRPs (based on the content of the HOME-instrument) and the monitoring and multidisciplinary consultation of DRPs in primary care.

**Objective**
This study aimed to develop eHOME, a mobile version of the HOME-instrument that includes a monitoring and consulting system for primary care.

**Methods**

**Design**
The development of eHOME was guided by the development phase of the framework for the development and evaluation of complex interventions of the Medical Research Council (MRC) [16]. The users and stakeholders were involved throughout the development phase by means of the human-centered design (HCD) for interactive systems [17] to develop a usable eHOME-instrument that fits the needs of the end users. The study was approved by the Medical Ethics Research Committee of the University Medical Center, Utrecht (the Netherlands) (11-129/C).

**Setting and Procedure**
This study was performed from September 2014 to March 2017 in a setting of home care teams, general practices, and pharmacies.

The HCD approach was divided into 6 phases: (1) the Delphi round; (2) the development, evaluation, and redesign of the prototype report webpage and the monitoring webpage interfaces for home care; (3) the usability evaluation pilot of the report webpage and the monitoring webpage interfaces for home care; (4) the expansion and development of the monitoring and consulting webpage interfaces for home care, GPs, and pharmacists; (5) the usability evaluation pilot of the report webpage and the monitoring and consulting webpage interfaces; and (6) the development of the final webpages interfaces. Table 1 shows how the 6 phases are mapped into the HCD described in ISO 92410-201 [17].

**Delphi Round**
In October 2014, a Delphi round was conducted in a workgroup of 13 participants (a postdoctoral researcher, a project manager, 2 pharmacists, 7 home care workers of one home care team, and 2 software developers). In the Delphi round, the context in which eHOME should be used and the requirements of the users for a report system (to report DRPs by home care workers who perform home visits), a monitoring and consulting system (for home care nurses, GPs, and pharmacists) and requirements on the organizational level were identified. During both parts of the round, the project manager played the role of an observer and reported the requirements on a whiteboard.
### Table 1. Methodological phases mapped into a human-centered design (HCD). The ✓ symbol shows which phases of this study belong to which phases of the HCD.

<table>
<thead>
<tr>
<th>Methodological phases</th>
<th>HCD(^a) phases</th>
<th>Specify the user requirements</th>
<th>Produce design solution to meet the user requirements</th>
<th>Evaluate the designs against requirements</th>
<th>Designed solution meets user requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Delphi round</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>The development, evaluation, and redesign of the prototype report webpage and the monitoring webpage interfaces for home care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation workgroup meeting</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Redesign report webpage and monitoring webpage interfaces for home care</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The usability evaluation pilot of the report webpage and the monitoring webpage interfaces for home care</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questionnaires</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semistructured interviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The expansion and development of the monitoring and consulting webpage interfaces for home care, GPs(^b), and pharmacists</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The usability evaluation pilot of the report webpage and the monitoring and consulting webpage interfaces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The development of the final webpage interfaces</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)HCD: human-centered design.

\(^b\)GPs: general practitioners.

### Development, Evaluation, and Redesign of the Prototype Report Webpage and the Monitoring Webpage Interfaces for Home Care

Two software developers developed a prototype of the report webpage and the monitoring webpage interfaces and a link between these 2 interfaces, based on the formulated requirements of phase one. In a second workgroup meeting, the prototypes of the webpage interfaces were presented and installed on the mobile phones and personal computers of the workgroup members, and they were tested by the home care workers. The report webpage interface was tested by the home health care assistants, who performed home visits, in a case study exercise. Several home care practice-based signs and symptoms of DRPs were reported by the report webpage interface in the exercise. The home care nurse evaluated the usability of the monitoring webpage interface for the home care nurses and assessed whether the reported problems of the home health care assistants were forwarded from the report webpage to this webpage interface. This process allowed the home care workers to get used to the webpage interfaces for the usability evaluation pilot. Additionally, the workgroup meeting resulted in a plenary discussion with a set of adjustments to improve the usability of the prototype webpage interfaces. By the adjustments as a result of the plenary discussion, the content of the signs and symptoms of the report webpage for the home care was redesigned. This phase was carried out from September to October 2014.

### Usability Evaluation Pilot of the Report Webpage and the Monitoring Webpage Interfaces for Home Care

To assess the usability of the webpage interfaces, a usability evaluation pilot among the home care team was performed. During a 6-week pilot period (November to December 2014), the home health care assistants (n=6), who performed home visits, reported signs and symptoms of DRPs by the report webpage interface. At the end of the pilot period, the home health care assistants answered a questionnaire about the usability of the report webpage interface. To explore the usability of the monitoring webpage, a semistructured interview with the home care nurse was carried out. The interview was audiotaped and transcribed verbatim. Data of the questionnaires and the semistructured interview resulted in a set of requirements to improve the usability of eHOME.

### Expansion and Development of the Monitoring and Consulting Webpage Interfaces for Home Care, General Practitioners, and Pharmacists

Between January and May 2015, the software developers expanded the monitoring webpage interface for the home care nurse with a consulting system. Furthermore, the monitoring
and consulting webpage interfaces for the GPs and the pharmacists and a link between the monitoring and consulting webpage interfaces that allows 2-way communication between the home care nurse and the GP or the pharmacist was developed.

**Usability Evaluation Pilot of the Report Webpage and the Monitoring and Consulting Webpage Interfaces**

To test the possibility of communication between the home care nurses and the GPs or the pharmacists through the monitoring and consulting webpage interfaces, a second evaluation pilot of 3 months was initiated. Using a convenience sample strategy, 2 groups were selected. The first group consisted of 12 home care teams and 3 pharmacies. The pilot of group one took place from November 2015 to February 2016. The second group consisted of 6 home care teams, 7 general practices, and 6 pharmacies. The pilot in the second group took place from March to June 2016. Before the start of the pilot, participants received information about the types and the consequences of DRPs in home care patients, the importance of early recognition of DRPs, the webpage interfaces, and an explanation of the pilot in a workshop meeting. The workshop was led by a home care nurse, 2 pharmacists, and the first author (ND). During both the pilots, the home care workers, who performed home visits, reported the signs and symptoms of DRPs by the report webpage interface, and the home care nurses performed daily triage of the reported signs and symptoms with the monitoring and consulting system. Nurses in group one could report problems to a pharmacist and nurses in group two could report problems to a GP and or a pharmacist by using the monitoring and consulting webpage interfaces. The GP or the pharmacist was asked to respond to the signs and symptoms by the monitoring and consulting webpage interfaces. Anonymized descriptive statistics of consultation data between home care nurses, GPs, and pharmacies showed whether the consultation between the users occurred. Semistructured interviews with 8 home care workers were performed following the pilot. The interviews were held at the home care workers’ home or the home care practice. Besides the interviews, every 2 weeks, an evaluation with the home care nurses (n=18) through a phone call was carried out. The interviews and the evaluation with the home care nurses aimed to explore the strengths and the weaknesses of the webpage interfaces and to explore the potential solutions. The interviews and biweekly evaluations were carried out by the home care nurse of the workgroup and ND. The interviews were audiotaped and transcribed verbatim.

**Development of the Final Webpages Interfaces**

Data of the semistructured interviews and the biweekly evaluations were used to adjust the webpage interfaces so that a final version of the webpage interfaces that meets the user requirements could be designed. The final version of the webpage interfaces was developed by the software developers between October 2016 and February 2017.

**Results**

The results are described for the following 6 phases: (1) the Delphi round; (2) the development, evaluation, and redesign of the prototype report webpage and the monitoring webpage interfaces for home care; (3) the usability evaluation pilot of the report webpage and the monitoring webpage interfaces for home care; (4) the expansion and development of the monitoring and consulting webpage interfaces for home care, GPs, and pharmacists; (5) the usability evaluation pilot of the report webpage and the monitoring and consulting webpage interfaces; and (6) the development of the final webpage interfaces.

**Delphi Round**

**Context of Use**

To achieve optimal management of DRPs by eHOME, it was determined that eHOME should be used by a multidisciplinary team of home care workers, GPs, and pharmacists. The goal for optimal management of signs and symptoms of DRPs in home care patients by eHOME is twofold. First, home health care assistants who visit clients for essential care (activities of daily living such as bathing) report the signs and symptoms. Second, a bachelor trained home care nurse performs a daily triage of the reported signs and symptoms and forwards problems to a GP or a pharmacist when their expertise is needed. Subsequently, the GP and pharmacist send feedback on the DRPs to the home care nurse. It was decided that eHOME must consist of 2 webpage interfaces: a report webpage interface for home health care assistants who perform home visits, and monitoring and consulting webpage interfaces for home care nurses, general practices, and pharmacies.

**User Requirements eHOME**

Requirements for both webpage interfaces and requirements on the organizational level were formulated.

**Report Webpage Interface**

The content of HOME-instrument [15] formed the basis for the report webpage interface. Home care workers unanimously decided that the report webpage interface needs to include the signs and symptoms of the HOME-instrument, 6 common home care practice-based problems: (1) medication used has not been listed on the medication list; (2) medication on the medication list is not in use; (3) home care patient uses other amounts or dosages; (4) home care patient uses medication on another time (notice which medicines and when); (5) thick legs or feet; and (6) wounds, and a possibility to add any other potential sign or symptom of a DRP. To monitor DRPs for a longer period (eg, thick legs or feet or bruises), the possibility to add photos of these observations was required.

**Monitoring and Consulting Webpage Interfaces**

The monitoring and consulting webpage interfaces must contain the reported DRPs of home care patients out of the report webpage interface (with information of client’s name, sex, and date of birth), possibility to add new profiles of home care patients with detailed identification details (eg, name[s], sex, date of birth, name of GP, and pharmacist).

**Organization Level**

A link between the report webpage and the monitoring and consulting webpage interfaces of the home care and a link between the monitoring and consulting webpage interfaces of
home care nurses and GPs and pharmacists was required so that the information of reported problems between the webpage interfaces could be shared. Furthermore, the webpage interfaces needed to be available for smartphones, tablets with the iOS and Android operating systems, as well as personal computers. All webpage interfaces must contain a log-in screen, so personal details of patients could be assessed by the users only by entering a username and a password in the log-in screen.

Development, Evaluation, and Redesign of the Prototype Report Webpage and the Monitoring Webpage Interfaces for Home Care

One report webpage interface was developed in which the DRPs could be reported and forwarded to the monitoring webpage of the home care nurse. For the home care nurse, one monitoring webpage was developed for incoming DRPs. All home care workers were able to report the problems of the case study exercise by the report webpage interface. Additionally, it was found that the reported problems appeared all in the monitoring webpage interface. Following the evaluation pilot, it was decided that the content of the first version of the report webpage (presented in Multimedia Appendix 1) needed to be divided into several webpages to improve the usability. Instead of one report webpage, including 33 signs and symptoms, 10 webpages were developed. The 33 signs and symptoms were replaced into 2 main categories (represented with an icon) and 7 subcategories. Furthermore, some textual changes were made to enhance the readability. Specification of the changes and the final content of the report webpage interfaces are presented in Multimedia Appendix 2.

Usability Evaluation Pilot of the Report Webpage and the Monitoring Webpage Interfaces for Home Care

Seven home care workers of the workgroup took part in the first pilot. Home health care assistants, who performed the home visits (n = 6), indicated that the report webpage interface is a convenient, clear, and a usable instrument. One home health care assistant missed the possibility to review their own reported observations, and 3 home health care assistants reported the need for a conformation that an observation was forwarded to the home care nurse. The home care nurse commented in the interview adjustments to optimize the monitoring webpage interface; first, a link between the electronic patient file and the monitoring and consulting webpage interface was mentioned so that patient data will automatically be added to the webpage interfaces. Thereby adding a client profile manually to the monitoring webpage can be avoided which is time-saving. Second, the possibility to forward observations of signs and symptoms to a GP and a pharmacist was desired so that phone-calls can be avoided, which is in turn time-saving.

Expansion and Development of the Monitoring and Consulting Webpage Interfaces for Home Care, General Practitioners, and Pharmacists

As a result of the usability evaluation pilot, it was possible to expand the monitoring webpage for the home care nurse with a consulting service and to develop a monitoring and consulting webpage for GPs and pharmacists and a link between the monitoring and consulting webpage of the home care nurse and the GPs and pharmacists.

Usability Evaluation Pilot of the Report Webpage and the Monitoring and Consulting Webpage Interfaces

Home health care assistants indicated in interviews that they were more aware of the problems related to medication use because of the signs and symptoms presented in the report webpage. Furthermore, the report webpage was considered to be easy to use, and specific problems of the various categories of medication problems and body symptoms were usable and easy to find. Home health care assistants were able to add a note and a photo, but only after a problem was already sent to the home care nurse. This sequence was perceived as not logical, and a reverse sequence was mentioned as a solution. The home care nurses were satisfied with the clear overview with the types and amount of problems per client of the monitoring and consulting webpage. Home care nurses indicated that by using the monitoring and consulting system service, more collaboration with pharmacists and GPs were experienced leading to more solutions for DRPs. Home care nurses informed home health care assistants about the feedback of a GP or pharmacist by a telephone call, which was experienced as time-consuming. The solution to forward feedback of a home care nurse, a GP, or a pharmacist from the monitoring and consulting webpage to the report webpage was mentioned. Furthermore, home care nurses indicated that the daily triage of the report problems was necessary but not performed as intended because of workload and because they needed to get used to the daily task. Therefore, some home care nurses decided to share the triage task with nurses who had the capability to perform a triage. The home care nurses indicated that if a problem is reported during a home visit, a notification by means of a pop-up on their mobile device will ensure that daily triage will take place. Adding new client profiles in the monitoring and consulting webpage was experienced as time-consuming. During the pilot, the home care nurses decided to share this task with home health care assistants. Home care workers mentioned that a link between the electronic patient file and the monitoring and consulting webpage is desirable and will ensure that data of clients who are added to the home care will be automatically added to the monitoring and consulting webpage. Extraction data showed that communication between the home care and pharmacies in group one and communication between home care and general practices in group two using the monitoring and consulting webpage interfaces was possible during the pilot.

Development of the Final Webpages Interfaces

In the report webpage interface, the sequence to add a note and a photo to a problem before sending the problem to the home care nurse was adjusted. Furthermore, this webpage was expanded with an extra webpage, for the feedback of home care nurses, GPs, and pharmacists on reported DRPs (see Figure 1 for screenshots of the report webpage interface). The following aspects of the monitoring and consulting webpage interfaces were modified: a consulting service between the monitoring and consulting webpage interfaces of nurses and the report webpage interface of home health care assistants; a pop-up
notification for incoming problems for home care nurses, general practices, and pharmacies; a link between the electronic patient files and the report webpage and the monitoring and consulting webpage for home care; and a link between the monitoring and consulting webpages of GPs and pharmacies to facilitate consultation between these disciplines (see Figure 2 for a screenshot of the monitoring and consulting web page interfaces).

**Figure 1.** Screenshot of the report web page interface.

**Figure 2.** Screenshot of the monitoring and consulting web page interfaces.
**Discussion**

**Principal Findings**

This study resulted in eHOME, a mobile version of the HOME-instrument and a monitoring and consulting system for primary care that will help home care workers to report observed DRPs during home visits and to communicate DRPs with GPs and pharmacists.

In our study, a systematic development of the eHOME webpages was conducted. The systematic development of this study reflects the phases of an HCD therefore, the methods of this development study are mapped into the HCD approach [17]. The HCD approach process enabled collaborative decision making by home health care assistants, home care nurses, GPs, and pharmacists for the design and development of eHOME webpage interfaces, which was found to be convenient, clear, and easy to use.

**Limitations**

Within the HCD approach, it is important to involve all humans (eg, users and stakeholders) from the beginning of the development process to develop a system that fits the context in which the system will be used and to meet the user requirements and technical requirements, which may increase the usability in clinical practice. In this study, the GPs, one group of users, are not involved in the context of use and user requirements analysis and the first usability evaluation pilot. Even though GPs were not involved in the decision-making process of the context of use and user requirements analysis and the first usability evaluation pilot, the second usability evaluation pilot showed that the monitoring and consulting webpage was usable for their clinical practice.

With the help of this study, the report and monitoring and consulting webpage interfaces have been developed however, the effectiveness of the multidisciplinary approach of DRPs by eHOME on patient outcomes is not yet known. Further research on the clinical effectiveness of eHOME on patient outcomes is needed.

**Comparison With Prior Work**

Several paper-based report tools for DRPs are available for home care patients [11-15]; however, to our knowledge, eHOME is the first digital tool for the report of DRPs in combination with a monitoring and multidisciplinary consultation service between the home care, general practices, and pharmacies.

Previously, other paper-based screening tools were converted to mobile versions, for example, D-VAS for pain assessments [18], MOST-92610 [19], the ACEmobile [20] for assessments of neurocognitive disorders, the CVD risk assessment app [21], and the Risk detection app (in Dutch: Risico signaleren app) [22]. However, detailed information regarding the methodology used to transit from the paper screening tool to a digital system has, to our knowledge, never been published. This information is of importance to determine which different phases of the transition process lead to the usability of a system and to learn from barriers and strengths of the transition process.

This study shows how through a systematic approach a paper tool was transformed to a digital system. Developing a paper-based digital system by a systematic approach is expected to enhance the usability in clinical practice. Furthermore, the transparency of this systematic process is of importance and helps others to plan and manage methodological phases of the transition process of paper-based tools into usable digital systems, when and how health care professionals and other stakeholders can be involved, and to consider barriers and strengths of a development process.

Within this study, the focus on recognition of DRPs was for older patients by the home care workers. However, in intramural care settings, such as nursing homes and hospitals, older patients are also vulnerable for DRPs and dependent on health care professionals. Therefore, eHOME can be used by health care professionals in several care settings.

**Conclusions**

By employing an HCD approach, the HOME-instrument was converted to eHOME webpage interfaces, which was considered convenient, clear, and easy to use for the report of the signs and symptoms of potential DRPs in home care patients and for the monitoring and multidisciplinary consultation of these problems in primary care. This study provides a description of a systematic approach that can be used for future development of digital systems of paper-based tools.

**Acknowledgments**

The authors gratefully acknowledge all home care workers, GPs, and pharmacists who participated in this study. The authors express gratitude to Judith van Leeuwen-Jans, Ineke Hammann-Borgdorff, Aouktje Kaspers, Eric Hiddink, and Ineke Zekveld who assisted during the Delphi round and usability evolution pilots. The authors appreciate the work of Code Yellow, who developed and redesigned the eHOME report webpage and monitoring and consulting webpage interfaces.

**Conflicts of Interest**

None declared.

**Multimedia Appendix 1**

Content of the first version of the report webpage for home care.
Multimedia Appendix 2

Content of the second version of the report webpage for home care.

References


Abbreviations

- DRPs: drug-related problems
- eHealth: electronic health
- eHOME: electronic HOME
- GPs: general practitioners
- HCD: human-centered design
- HOME-instrument: home care observation of medication-related problems by home care employees instrument
- MRC: Medical Research Council
Human Factors and Data Logging Processes With the Use of Advanced Technology for Adults With Type 1 Diabetes: Systematic Integrative Review

Marion Waite1, MSc; Clare Martin2, DPhil (Oxon); Rachel Franklin3, MSc; David Duce2, PhD; Rachel Harrison2, PhD

Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, United Kingdom
Faculty of Technology, Design & Engineering, Oxford Brookes University, Oxford, United Kingdom
Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Marion Waite, MSc
Faculty of Health and Life Sciences
Oxford Brookes University
Marston Rd Campus, Jack Straws Lane
Headington
Oxford, OX3 7AB
United Kingdom
Phone: 44 07500619071
Email: mwaite@brookes.ac.uk

Abstract

Background: People with type 1 diabetes (T1D) undertake self-management to prevent short and long-term complications. Advanced technology potentially supports such activities but requires consideration of psychological and behavioral constructs and usability issues. Economic factors and health care provider capacity influence access and uptake of advanced technology. Previous reviews have focused upon clinical outcomes or were descriptive or have synthesized studies on adults with those on children and young people where human factors are different.

Objective: This review described and examined the relationship between human factors and adherence with technology for data logging processes in adults with T1D.

Methods: A systematic literature search was undertaken by using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Quality appraisal was undertaken and data were abstracted and categorized into the themes that underpinned the human factor constructs that were examined.

Results: A total of 18 studies were included. A total of 6 constructs emerged from the data analysis: the relationship between adherence to data logging and measurable outcomes; satisfaction with the transition to advanced technology for self-management; use of advanced technology and time spent on diabetes-related activities; strategies to mediate the complexities of diabetes and the use of advanced technology; cognition in the wild; and meanings, views, and perspectives from the users of technology.

Conclusions: Increased treatment satisfaction was found on transition from traditional to advanced technology use—insulin pump and continuous glucose monitoring (CGM); the most significant factor was when blood glucose levels were consistently <7.00 mmol/L ($P \leq 0.01$). Participants spent considerable time on their diabetes self-care. Logging of data was positively correlated with increasing age when using an app that provided meaningful feedback (regression coefficient=55.8 recordings/year; $P \leq 0.01$). There were benefits of CGM for older people in mediating complexities and fears of hypoglycemia with significant differences in well-being ($P \leq 0.001$). Qualitative studies explored the contextual use and uptake of technology. The results suggested frustrations with CGM, continuous subcutaneous insulin infusion, calibration of devices, and alarms. Furthermore implications for “body image” and the way in which “significant others” impacted on the behavior and attitude of the individual toward technology use. There were wide variations in the normal use of and interaction with technology across a continuum of sociocultural contexts, which has implications for the way in which future technologies should be designed. Quantitative studies were limited by small sample sizes, making it difficult to generalize findings to other contexts. This was further limited by a sample that was predominantly white, well-controlled, and engaged with self-care. The use of critical appraisal frameworks demonstrated where research into
human factors and data logging processes of individuals could be improved. This included engaging people in the design of the technology, especially hard-to-reach or marginalized groups.

(JMIR Hum Factors 2018;5(1):e11) doi:10.2196/humanfactors.9049

KEYWORDS

adult; type 1 diabetes mellitus; T1D; technology; self-management; self-care; technology; telehealth; telemedicine; reminder system; continuous glucose monitoring; Sensor-augmented pump therapy; closed loop systems; adherence; compliance; barrier; usability

Introduction

Personal decision-making and human factors

Individuals with type 1 diabetes (T1D) are confronted with complex tasks through which to manage their blood glucose (BG) levels. T1D is an autoimmune disease where the beta cells in the pancreas no longer produce insulin, resulting in dangerously high BG levels or hyperglycemia. The person diagnosed with T1D is subsequently required to self-administer insulin. This involves regular self-monitoring of BG levels and calculation of appropriate insulin doses. There is a delicate balance between the reductions of the risks of long-term complications (often associated with hyperglycemia) and those of hypoglycemic events. This puts emphasis on adherence and patient behaviors. It has been suggested that large numbers of people with T1D are nonadherent [1]. Additionally, Patton [2] highlights multiple social, emotional, and cognitive barriers. The prevalence of new and emergent technologies to support self-management of T1D through personal data logging processes and support for decision making may have the potential to address these issues.

There may be a dilemma for health care providers due to the economic implications of adopting such technologies for individuals compared with potential public health benefits. This raises the issue of identification of adults with T1D who may benefit the most. There are associated questions around how to investigate and evaluate the benefits of such technology with respect to specific populations in such a way as to inform future design decisions. Thus, consideration of psychological and behavioral constructs alongside evaluation of the usability of devices, also known as human factors, is an integral component of any investigation that involves clinical consideration for emergent technology aimed at self-management of T1D.

The objective of this review was to describe the relationship between human factors and technology adherence for data logging processes in adults with T1D and to explore the factors that influence this association.

Background

Advanced Technology for Self-Management of Type 1 Diabetes

The potential for technology to support individuals with T1D is increasing rapidly. The following overview covers general principles where the individual interacts with the technology to log his or her personal data in some capacity.

Continuous glucose monitoring (CGM) provides information regarding changes in glucose concentrations within interstitial fluid in real time. The corresponding device consists of a sensor, which is placed in the subcutaneous tissue, and a monitor, which may or may not be connected wireless. CGM data are used either to assist with retrospective decision making by a clinician or to support individual self-management. There is potential for an abundance of information about trends and directions in BG levels, including fluctuations over time for retrospective analysis [3]. One of the motivations for development of CGM is to recognize nocturnal hypoglycemia; another is to support people who may have lost their hypoglycemic awareness [4].

Real-time CGM has been available from 2005, and since then, advances in technology have improved the accuracy of CGM systems and provide potential advantages in terms of relaying the glucose history of an individual. Castle and Jacobs [5] suggest that there is valid evidence that both hyperglycemia and hypoglycemia are reduced with consistent CGM use. The optimal way to adjust insulin doses is complex, and there is little guidance for individuals about how to interpret the data. Internationally, there is low uptake of CGM but that may say more about availability and access than about the wishes of individuals.

Most individuals with T1D administer insulin via multiple daily injections (MDI), but some use an insulin pump that delivers bolus doses of insulin on demand of the user in addition to tiny amounts of insulin. These are administered every few minutes but may vary at different times of the day, thereby delivering what is known as continuous subcutaneous infusion of insulin (CSI). Advantages may include not physically injecting each delivered bolus dose and the availability of more physiologic basal insulin than available long-acting insulins can provide, and it is not necessary to inject each time a dose is administered. Theoretically, the way in which doses may be tailored is more specific to the insulin requirements of the individual [3]. There are 2 types of insulin pumps. One is tethered to a cannula that enters the subcutaneous tissues. This means that the pump must be worn by the user and may be visible. A patch pump on the other hand consists of a short tube attached to a cannula with an integrated micropump that is controlled wirelessly by the user [3], which can be hidden.

Sensor-augmented pump therapy (SAPT) is the concurrent application of real-time CGM with an insulin pump. However, this does not lead to automatic insulin adjustment. It is incumbent on the user to use adjunctive self-monitoring of BG and make dose adjustments to suit his or her own insulin requirements. Future developments include decision-support systems that will recommend insulin doses based on an array

http://humanfactors.jmir.org/2018/1/e11/
of factors, including historical data of the individual, and will also connect to health care providers.

Closed loop systems are sometimes known as artificial pancreas and manage insulin delivery in response to real-time CGM data, which is controlled by algorithms rather than preprogrammed rates [6]. According to Castle and Jacobs [5], this can also include delivery of glucagon to raise BG levels when necessary.

Apps run on mobile devices such as mobile phones and tablets and perform functions previously restricted to personal computers. Those designed specifically for people with T1D can generally be categorized into 5 areas:

1. Glucose tracking diaries
2. Carbohydrate estimators
3. Recipe planners
4. Medication adherence tools
5. Diabetes education platforms [7]

Telehealth refers to logging of health care data by the patient, which is tracked by health care professionals (HCPs) at a distance [8]. For example, the use of mobile devices by the patient enables any time, any place, anywhere logging and transmission of data.

Access, Uptake, and Current Limitations

Access and uptake of advanced technology, such as CGM and CSII, are controlled by health care economies and clinical policy guidelines. For example, in 2011, it was estimated that uptake may be between 20% and 30% in the United States and Israel compared with 1% in Denmark [9].

Acerini [10] claims that, even if CGM and CSII were readily available, those who could benefit the most from use would not access it and that diabetes technology uptake is lower in some ethnic groups. Furthermore, adoption is governed by socioeconomic status and cultural factors in addition to access to appropriate health care services. Crucially, health care practitioners’ willingness and capacity to support patient access are other critical factors [11].

To date, most research into use of advanced technology has focused on the clinical outcomes, which overall are equivocal [9,10,12]. Kerr and Partridge [6] critique the endpoints of previous clinical trials, which focus purely on glycated hemoglobin (HbA1c) levels without reference to other outcomes that may be equally meaningful to adults with T1D.

Transition and use of advanced technologies require training and physical and psychological adaptation by the users and their families. Human factors are, therefore, an essential component in reaching a better understanding of uptake and use of technology and in informing design decisions.

Human Factors and Type 1 Diabetes

There are differential aspects of the human factor that affect the use of technology in diabetes self-management [13]. These may be conceptualized as follows:

1. Behavioral
   • Barriers to adherence [2]
2. Psychological
   • Adjustment to diabetes [15]
   • Fear of hypoglycemia [11,14,16]
   • The emotional implications of increased responsibility for self-management including fear of disapproval by HCPs and worthiness to receive cutting-edge treatment [17]
   • Self-belief, impact on quality of life, reactions of others, unconscious motives based on earlier experiences [18]
   • Trust in the technology, letting go of prior routines [11,17]
   • Depression and eating disorders [18]
3. Social
   • Wearability of devices and body image [11]
   • Interpersonal relationships and working out how to handle interactions with others and when and how to disclose the condition [18]
   • Support from significant others to engage with technology [9]
   • Choice about whom to share data with [11]
   • Stigma surrounding the carrying out of tasks in social situations [4]
4. Cognitive
   • Educational needs, such as that of learning how to use the technology and utilize greater knowledge of personal glucose trends to make dosing decisions [9]
   • Additional learning associated with the use of technology [19]
   • Health literacy and associated embarrassment with low literacies [20]
   • Reduced cognitive abilities associated with age and adult level of educational attainment [21]

Current research in the field of advanced technology for diabetes has emerged from different disciplines, for example, health care practice, psychology, computer science, electronic engineering, and related industries. To reach a full understanding, it is crucial to bring this research together in a systematic way. Previous reviews have focused on clinical outcomes alone [5,22], have descriptively scoped the literature [13,23], or have synthesized studies on children and young people with studies on adults [24] where the needs for technology and associated human factors are likely to be different. Thus, there is a gap for a review that systematically appraises current research on the relationship between human factors and data logging processes with advanced technologies for adults with T1D.

Aims of the Review

The aim of this systematic review was to describe the relationship between human factors and adherence with technology for data logging processes in adults with T1D and to explore the factors that influence this association.
An integrative literature review research design was chosen because it provides a more holistic conceptualization on a complex topic [25] such as human behavior and facilitates inclusion of diverse methodologies and theories, given the interdisciplinary approach toward research in the field.

A protocol was developed (Multimedia Appendix 1) to clarify the aims, sampling strategy, exclusion and inclusion criteria, methods, outcomes, language, and search strategy.

**Methods**

**Literature Search**

A systematic search of the literature was performed in accordance with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) [26] in January to March of 2017 (Multimedia Appendix 2).

The following databases were searched: Computing Research Repository (2006 to January 2017); PsycINFO, EMBASE, and MEDLINE (2006 to January 2017); Web of Science (2006 to January 2017); Zetoc (2006 to January 2017); Excerpta Medica and Scopus (2006 to January 2017); and ProQuest (2006 to January 2017). Only research that was undertaken during the last 10 years was included as technology for the self-management of T1D has been developing rapidly during this time. Search terms included: Diabet* AND Techno* AND Behavi*; Self-manage* OR self-manage* OR manage* OR self-care OR self-care; technolog* OR telehealth OR telemedic* OR reminder system* OR text messag* OR application OR app*; adhere* OR compliance OR barrier OR problem* OR obstacle; MH Diabetes Mellitus, Type 1*.

Searches were limited to adults (over 18 years) and filtered to studies of adults published in English. Reference lists were also searched in addition to subject-specific websites and key journals (Multimedia Appendix 1). The search strategy was carried out in collaboration with a university health care librarian. Unpublished studies (dissertations and theses) were excluded, in addition to editorials, opinions, and discussion papers. Studies were reviewed for the following criteria: (1) primary research; (2) empirical data on adherence to data logging processes with the use of advanced technology for adults with T1D; (3) an investigation of the relationship with psychological, social, and human factors; and 4) the psychological outcome measures were explicit (quantitative studies) or alternatively included a clearly described picture of the phenomenon that included the user perspective (qualitative studies).

**Search Outcomes**

The search strategy produced 1 article in the Computing Research Repository; 348 articles in PsycINFO, EMBASE, and MEDLINE; 40 articles in the Web of Science; 84 articles in Zetoc; 38 articles in Excerpta Medica and Scopus; and 36 articles in ProQuest. Once duplicates were removed, additional articles were excluded due to limitations associated with unclear abstracts or for not meeting the inclusion criteria (ie, children, type 2 diabetes, and gestational diabetes). In total, 72 citations were retained and each abstract was read for relevance.

Also, 3 citations were found from searching reference lists and key journals. One study, which included children and their carers, was retained because outcomes were compared with adults who also participated within the study [27]. To reduce bias and ensure that only the most relevant articles were selected, the second and third authors reviewed the titles and abstracts according to the protocol criteria and a consensus was reached about the articles to be included in the review. In total, 22 articles met the inclusion criteria, and these included 14 quantitative studies, 5 qualitative studies, and 3 mixed-method studies (Multimedia Appendix 2).

**Quality Appraisal**

Whittemore and Knaff’s approach [25] of using as many instruments as necessary to evaluate the quality of the data was taken because this is an integrative review, and the data are drawn from more than one disciplinary area that use a range of research traditions that align with quantitative, qualitative, or mixed-method research designs. The instruments for appraisal were selected from the University of South Australia International Centre for Allied Health Evidence [28] databases of critical appraisal tools. The following criteria were taken into consideration for types of study design: demographic information of the participants and statement of research question, appropriateness of the research question for the selected study design, and approach to recruitment reported (Table 1). The criteria for quantitative study designs included power analysis reported response rate, reliability and validity of study instruments and method of data analysis (Table 2). The following criteria were considered for qualitative studies: theoretical perspectives, audit trail, member checks, peer review of qualitative data, and method of data analysis (Table 3). The first author undertook the quality appraisal of each study, which was peer-reviewed independently by second and third authors. Following the critical appraisal process, 4 studies were further excluded for poor methodological design.

**Data Abstraction and Analysis**

The review data were categorized and synthesized into the themes that underpinned the human constructs that were examined and the outcomes that were reported. Mile and Huberman’s [46] approach to coding of data, which involves data reduction and comparison, was utilized.
Table 1. Quality appraisal.

<table>
<thead>
<tr>
<th>Author</th>
<th>Type of study design</th>
<th>Aptness of study design for research aims</th>
<th>Demographic information of participants</th>
<th>Approach to recruitment reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groat et al [29]</td>
<td>Observational</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Gonder-Frederick et al [30]</td>
<td>Observational</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Skrosveth et al [31]</td>
<td>Observational</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Tansey et al [27]</td>
<td>Randomized controlled trial</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Kamble et al [32]</td>
<td>Randomized controlled trial</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Martinez- Sarrigui et al [33]</td>
<td>Randomized controlled trial</td>
<td>Yes</td>
<td>Yes</td>
<td>Not reported</td>
</tr>
<tr>
<td>Gonzalez- Molero et al [34]</td>
<td>Longitudinal cohort study</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Kirwan et al [35]</td>
<td>Randomized controlled trial</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Polonsky et al [36]</td>
<td>Cross-sectional</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Barnard et al [37]</td>
<td>Cross-sectional</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Naranjo et al [38]</td>
<td>Cross-sectional</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Borges and Kubiak [39]</td>
<td>Cross-sectional</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Shepherd et al [40]</td>
<td>Qualitative</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ritholz et al [41]</td>
<td>Qualitative</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>O’Kane et al [42]</td>
<td>Ethnography</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Storni [43]</td>
<td>Ethnography</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lawton et al [44]</td>
<td>Qualitative</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Barnard et al [45]</td>
<td>Mixed methods</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2. Quality appraisal quantitative studies.

<table>
<thead>
<tr>
<th>Author</th>
<th>Power calculation reported</th>
<th>Response rate (%)</th>
<th>Reliability and validity of study instrument established</th>
<th>Method of data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groat et al [29]</td>
<td>No</td>
<td>Not reported</td>
<td>No</td>
<td>Correlation analysis</td>
</tr>
<tr>
<td>Gonder-Frederick et al [30]</td>
<td>No</td>
<td>Not reported</td>
<td>yc</td>
<td>Analysis of covariance</td>
</tr>
<tr>
<td>Skrosveth et al [31]</td>
<td>No</td>
<td>Not reported</td>
<td>No</td>
<td>Regression analysis</td>
</tr>
<tr>
<td>Tansey et al [27]</td>
<td>No</td>
<td>Not reported</td>
<td>Yes</td>
<td>Correlation analysis</td>
</tr>
<tr>
<td>Kamble et al [32]</td>
<td>No</td>
<td>Not reported</td>
<td>Yes</td>
<td>Correlation analysis</td>
</tr>
<tr>
<td>Martinez- Sarrigui et al [33]</td>
<td>No</td>
<td>Not reported</td>
<td>No</td>
<td>Regression analysis</td>
</tr>
<tr>
<td>Gonzalez-Molero et al [34]</td>
<td>No</td>
<td>Not reported</td>
<td>Yes</td>
<td>Correlation analysis</td>
</tr>
<tr>
<td>Kirwan et al [35]</td>
<td>Yes</td>
<td>Not reported</td>
<td>Yes</td>
<td>Regression analysis</td>
</tr>
<tr>
<td>Polonsky et al [36]</td>
<td>No</td>
<td>48.6</td>
<td>No</td>
<td>Regression analysis</td>
</tr>
<tr>
<td>Barnard et al [37]</td>
<td>No</td>
<td>96.7</td>
<td>No</td>
<td>Correlation analysis</td>
</tr>
<tr>
<td>Naranjo et al [38]</td>
<td>No</td>
<td>Not reported</td>
<td>Yes</td>
<td>Correlation analysis</td>
</tr>
<tr>
<td>Borges and Kubiak [39]</td>
<td>No</td>
<td>Not reported</td>
<td>Yes</td>
<td>Factor analysis</td>
</tr>
</tbody>
</table>
The 18 studies included in this review consist of 5 qualitative studies [40-44], 5 experimental studies [27,32-35], 3 observational studies [29-31], 4 cross-sectional studies [36-39], and 1 mixed-methods study [45]. Of the studies, 5 were smaller samples drawn from parent clinical trials [30,32,38,40,41].

The total number of participants who were included in the 18 studies was 3320 and the mean age was 42 years, although one study [36] specifically recruited people over the age of 65 years. Female participants represented 53% of the sample. The mean prebaseline HbA1c was 7.9% (where reported).

Multimedia Appendix 3 summarizes the type of technology included in the review and the human factor constructs and outcomes that were examined.

After categorization and synthesis of themes, 6 overall constructs emerged:

1. The relationship between adherence to data logging and measurable outcomes
2. Satisfaction with the transition to advanced technology for self-management
3. Use of advanced technology and time spent on diabetes-related activities
4. Strategies to mediate the complexities of diabetes and the use of advanced technology
5. Cognition in the wild
6. Meanings, views, and perspectives from the users of technology

The Relationship Between Adherence to Data Logging and Measurable Outcomes

There was inconclusive evidence about the relationship between adherence to data logging process and measurable outcomes. For example, Kirwan et al [35] examined a freely available iOS app—Glucose Buddy—combined with text messaging feedback from a diabetes educator aimed at the improvement of glycemic control. The intervention group showed a significant decrease in HbA1c (mean −1.10; SD 0.74; (P ≤.001) over the 9-month period of the study; however, linear regression showed no significant relationship between the level of engagement with the app and these outcomes. This result may be interpreted with caution, given the small sample size (n=27). Furthermore, there was a potential socioeconomic bias in that participants were required to have iOS ownership.

Groat et al [29] analyzed individual participant internet protocol address data to characterize the relationship between adherence to insulin bolus dosing, logging of carbohydrate intake, and BG monitoring and glycemic control for a 1-month period. The only significant outcome was that an increase in daily insulin bolus doses had an impact on increasing the number of days that the BG was at target (r=.93). The reported results were based upon an extremely small sample (n=8) and described as regression analysis, which contradicts the researchers’ claims for undertaking a qualitative study.

Satisfaction With Transition to the Use of Advanced Technology for Self-Management

Some findings suggest that adults with T1D may feel more satisfied with their treatment on transition to advanced technology. For example, Gonzalez et al [34] evaluated the overall effect of adding a telemedicine system for adults with T1D who were treated with an insulin pump and real-time CGM. This was a longitudinal study that measured the physical and psychological outcomes of the intervention. Mean plasma HbA1c was significantly lower at 6 months compared with prebaseline (6.97 vs 7.5; P=.01); there was a significant reduction in glucose variability at 6 months compared with baseline (53.1 vs 68.7; P=.04) and prebaseline (53.1 vs 67.3; P=.04), and time spent interacting with the sensor correlated positively with time in normoglycemia (r=.72; P=.03) and negatively with occurrences of mild hypoglycemia (r=.64; P=.02). From a psychological perspective, there was an improvement in quality of life scores at 6 months in comparison with baseline (92.4 vs 86.9; P=.01), and participants with poorer glycemic control had significant improvements with prior dissatisfaction with treatment (34.3 vs 31.6; P=.01).

The authors acknowledged that the findings were based on a small sample size (n=15), and therefore, it is not possible to generalize the outcomes. The authors also questioned whether the point of being observed affected the outcome measures. However, the study did show that there may be benefits for well-controlled individuals using CGM in conjunction with telemetry in terms of HbA1c and quality of life as reported in the previous paragraph.

There is some consistency regarding the perceived physical outcomes and satisfaction of the above study with the findings of Barnard et al [45], who measured the relationship between satisfaction when transitioning to the then-current insulin pumps (Animas Vibe CGM-enabled system IV) and personal glycemic control. The most significant contributing factor to treatment satisfaction was when BG levels were consistently <7.00
mmol/L ($P=0.009$). The limits of this study are that the findings are based on self-report, and it is not clear why only 22 items of the 50 on the Insulin treatment satisfaction questionnaire were included on the survey instrument.

**Use of Advanced Technology and the Relationship With Time Spent on Device-Related Activities**

Frequent users of existing diabetes technology may find an easier transition to more advanced options. For example, Tansey et al [27] examined the perceived barriers and benefits to CGM use and how this related to frequency of use. Engaged CGM users were more satisfied, with higher frequency users less bothered by the “hassles” of the device. Frequent users were classified as engaged with CGM for more than 6 days per week and infrequent users less than 4 days per week. Adults and parents of users had higher total and subscale scores on the CGM satisfaction survey ($P=0.0009$). All respondents reported that visualization of glucose trends and the opportunity to detect hypoglycemia were the best aspects of use of CGM (text item responses in the questionnaire).

Martinez-Sarriegui et al [33] analyzed patient behavior when using the intervention of telemedicine system combined with CGM to identify how the CGM data captured participant interactions with the mobile system. In 2 phases of the experiment (with and without the telemedicine system), participants were provided with tools for visualization, management of monitoring data, and wireless downloading of data from an insulin pump via a personal smart assistant running on a personal digital device. The number of times interacting with the system was higher during the intervention phase (29.0 vs 18.8; $P=0.04$), and the total time spent interacting with the system was also higher during the intervention phase (04:27:11 vs 01:47:07; $P=0.009$).

Kamble et al [32] compared weekly estimates of time, changes in time, and patient time costs associated with diabetes-related care between SAPT and MDI. They used data on patient-reported time collected over a 52-week period. Participants were required to log the total time spent per week on diabetes management for a range of diabetes-related variables. The total time spent on the SAPT arm of the study was higher than time spent on MDI during and after pump initiation within the overall 52-week study. The reported weekly time estimates were as follows: SAPT 4.4 hours and MDI 3.4 hours (95% CI 0.4-1.7). However, all adults with T1D in the study reported that they spend considerable time on diabetes care.

Each of the above 3 studies suggests that engagement with technology is time consuming. Given that the inclusion criteria for the Tansey et al [27] study were prior high frequency of self-monitoring, it is not clear if the technology was a mediating factor for engagement. The Martinez-Sarriegui et al [33] study was limited by a small sample size (n=10) and did not include any details about how the study instrument was developed or how the participants valued the feedback from the telemedicine system. Furthermore, there was a possibility for margin of error with the Kamble et al [32] study as it was not clear how participants measured time costs.

**Strategies to Mediate the Complexities of Diabetes and the Use of Advanced Technology**

Some researchers have attempted to understand the way in which the human complexities of diabetes have the potential to be mediated with the use of advanced technology.

**Meaningful Feedback for the User**

Skrosveth et al [31] explored which methods of diabetes data analysis could be realistically used to provide meaningful feedback for the user. A mobile diary app was developed for adults with T1D to log insulin doses and dietary intake with options for the user to comment upon these and a screen to visualize each of the following variables: BG level, insulin dosing, and dietary intake. Retrospectively, the sample was divided into 2 groups: “adopters” (n=18), who reliably logged data for at least 80 days, and nonadopters (n=12), who did not. Logging of data was positively correlated with increasing age (regression coefficient=55.8 recordings per year; $P \leq 0.07$), but the usage did not significantly correlate with prestudy HbA$_1c$ ($P=0.33$) or gender ($P=0.09$). The researchers also found that several methods of pattern recognition were unable to predict future BG values. The study was limited by lack of demographic information about the participants and how they were recruited. More information about nonadopters such as confounding variables would have increased the reliability and validity of the results.

**Engaging Older Adults With Continuous Glucose Monitoring**

Polonsky et al [36] surveyed 2 groups of participants aged 65 years and older in T1D to determine differential characteristics between users of real-time CGM and nonusers (hopefuls). CGM hopefuls reported a higher incidence of 1 moderate hypoglycemic episode in the preceding 6 months (90% vs 78%; $P=0.04$). 1 hypoglycemic-related emergency room visit during the preceding 6 months (18.7% vs 6.7%; $P=0.002$), and 1 hypoglycemic event requiring assistance by another during the preceding 6 months (80% vs 57.6%; $P \leq 0.001$). CGM hopefuls also reported significant differences in well-being ($P=0.009$), hypoglycemic distress ($P=0.04$), and feeling of powerlessness ($P=0.04$). The study suggested potential benefits for older people with the use of advanced technology, which is important given that hypoglycemic unawareness increase with age. A drawback of the study was that the 2 groups were of unequal sizes: the user group=11 and the hopeful group=75.

**Information Overload and Ease of Use**

Borges and Kubiak [39] explored the relationship between information overload, ease of use, and personal attitude in the use of CGM by identification of motivations to use CGM and comparison of characteristics between groups with differing levels of CGM experience. The findings were that, irrespective of the level of experience, the advantages of CGM were perceived as high and the disadvantages perceived as low. There was a significant difference with respect to perceived information overload; adults with T1D without experience rated this higher than adults with T1D with more experience (90% CI 1.443-0.785; $P \leq 0.001$). This is important because information overload had a negative influence on the ease of use ($P \leq 0.001$).
The study reports statistically significant outcomes; however, the participants were recruited through Web-based forums and social media and described as having high levels education, which was a potential socioeconomic limiting factor.

**The Potential of Continuous Glucose Monitoring to Explore Stressors**

Gonder-Frederick et al [30] investigated the relationship between routine daily stressors, BG levels, and diabetes management strategies in a naturalistic setting using a CGM data to generate BG profiles (adults with T1D were also participating in multicenter cross-over randomized controlled trial closed-loop control CLC study). There was no relationship found between stress ratings and averages daily glucose. However, stress ratings were positively related to low BG levels ($P=0.025$). Overall, the results suggested individual differences between stress and glycemic control for people with T1D and the potential of CGM to explore this more in depth. This needs to be countered with the acknowledged small sample of participants ($n=33$) and a short-term study with highly selected participants.

**The Relationship Between Diabetes Distress and Technology**

Naranjo et al [38] undertook a comparative analysis of the level of diabetes distress that is associated with diabetes devices and technology between users of traditional technology (BG meters and MDI) and advanced technology (pump therapy and CGM). The results showed significant differences between attitudes to technology with CGM users being more positive than nonusers (24.87 vs 23.87; $P \leq 0.001$). Pump users were more positive than MDI users (24.8 vs 22.98; $P \leq 0.001$). There were no significant differences in distress across all types of technology use by participants. However, there was no account for confounding variables other than age.

Ritholz et al [41] qualitatively compared psychosocial differences between 3 groups of participants who were participants from the Juvenile Diabetes Research Trial: respondents ($n=7$), drawn from a primary cohort who had shown improvement in glycemic control; responders ($n=6$), drawn from a secondary cohort who had demonstrated a reduction in HbA1c in within target range, and nonresponders ($n=7$), who had a less than 0.5% reduction in HbA1c. The following themes emerged from the findings: coping with frustrations, use of CGM information, significant other information, and body image. Frustrations were experienced with CGM, CSII, calibrations, and alarms. Responders reported a self-controlling coping style whereas nonresponders were more likely to make an emotional response. All participants were engaged with minute-to-minute information, but responders were more likely to use retrospective information to spot trends and act upon them. Many responders reported significant other involvement, especially males who suggested that they did not allow other important agents about the risks of hypoglycemia. Body image of use of the device was associated with “nonresponders,” who felt uncomfortable about using the device in public places and intimate situations. The researchers raised the role of “significant others” in CGM research and suggested that this is an underexplored area. The research also highlights the clinical implications of preparation of adults with T1D to deal with frustrations and cognitive overload.

The limitations of the research are that it was carried out on a population that was described as well educated and homogenous.

**Cognition in the Wild**

Some researchers have adopted an ethnographic approach to explore how technology is used in the context of the everyday lives of adults with T1D.

O’Kane et al [42] took a sociocultural perspective and reported on 3 qualitative studies that examined how devices for adults with T1D are adopted, carried, and used in a variety of everyday contexts. This is based on the premise that adults with T1D are encouraged to self-regulate by HCPs, but the nature of everyday life is contingent upon the dynamics of the unfolding situation. The following themes emerged from the data analysis: misuse, inappropriate use, and unintended use of the technology. The authors’ main point is that any individual can report a wide variation in normal use of their technology across a continuum of public use, work-life use, and in the company of friends and family. This was based on the perceived emotions and attitudes of the other party within a given context. Uncertainty in discrete situations can lead to hiding a device, whereas showing off the device in other situations can lead to normalization and control of a situation. This corresponds with the findings of Ritholz et al [41], which were reported in the previous section regarding the place of significant others in uptake and use of technology.

**Meanings, Views, and Perspectives From the Users of Advanced Technology**

Research that examines the meaningfulness and perspectives of the user has an important role to play in the future and ongoing development of advanced technology. Shepherd et al [40] explored both desires and concerns regarding the use of CGM for self-management. The findings suggested that adults with T1D who already used insulin pumps and CGM had a diverse range of attitudes and concerns along a continuum regarding personalized glucose advisory systems. Participants would have liked advice from the system on suggestions for correction boluses, basal rates, insulin-carbohydrate ratios, and alerts to the risks of hypoglycemia. However, it would be necessary for the individual to understand how the advice was generated, trusting that all personal variables would be considered to develop the confidence to relinquish control to an automated system. A shortcoming of the study is that it was not entirely clear how the themes were arrived at.

Lawton et al [44] (2014) found evidence of similar themes during a longitudinal study of the use of insulin bolus calculators following the intervention of a dose adjustment for normal eating course. Adults with T1D were motivated by the device because it saved time and effort in calculations; however, those who were confident in their mathematical ability undertook their own individual calculations and were paradoxically less likely to use the device over time. Reliance on the calculator alone had a detrimental impact on glycemic control. Some participants left the ratios unchanged until their next clinician/study review, and for some, this was attributed to not knowing how to change the settings. Underconfidence in
carrying out personal calculations or not knowing how to change settings led to loss of trust in the technology.

Storni [43] contends that diabetes is more than a disease and should be regarded as a complex lifestyle. People with T1D develop lay expertise that is unique to their situation. This creates implications for technology design, and it is crucial to involve the user in the process. This perspective is based upon findings that emerged from an ethnographic study on diabetes support groups and by following individuals with T1D within the context of their everyday lives [43]. The purpose was to examine what participants really did in dealing with their condition as opposed to what they were told to do by clinicians. These findings influenced the design of a tagging system for events from everyday life to link them to carbohydrate intake and BG readings to create meaning between the events and a log for the individual on a mobile device. A shortcoming of the study is that the report provided a lack of demographic information about participants, which is important in qualitative research to determine transferability to other contexts. Nevertheless, there is an emergent field of research that addresses the diverse needs of people with T1D in the design of technologies.

Discussion

Principal Findings

Advanced technology for the management of T1D needs to have clear benefits that are meaningful to adults with T1D. The aim of the review was to describe the relationship between human factors and adherence with technology for data logging processes in adults with T1D and to explore the factors that influence this association.

There was inconclusive evidence about the relationship between adherence to data logging and measurable outcomes in relation to the review question. However, clinical values may have less importance than perceived outcomes for individuals. The review did suggest increased satisfaction with treatment on transition to advanced technology; however, this was biased toward frequent users of existing technologies and with an acceptance of the time required to spend on diabetes care.

The review also showed some benefits of advanced technology for older people by mediating complexities and fear of hypoglycemia. There appears to be a wide variation in the normal use of technology for adults with T1D across a continuum of sociocultural contexts. There is also a variability regarding user involvement in the design of future technologies and the role of “significant others” and this requires further research. People need to be able to trust technology as the capacity for intelligent decision-making advances.

In the literature that was reviewed, participants appeared to be a highly selective group biased toward white populations. Another limitation was the relatively small sample sizes of some of the quantitative studies included within the review, only 1 study [35] reporting on a power calculation, thus making it hard to generalize the findings.

A significant issue was that where demographic characteristics were reported (Table 1), 95% of the participants were described as white. The data suggest that those from higher socioeconomic groups are more likely to have access to and engage with technology in their self-management behaviors [38]. Of the studies, 2 [30,45] purposefully selected participants with prior adherent behaviors; however, 1 study recruited participants who were less engaged with technology and adherence [32].

The predominance of white participants, combined with the fact that 6 of the reviewed studies were samples drawn from parent clinical trials, suggests that the data are based on a highly selective group. This may not be representative of the general adult population with T1D. The mean baseline HbA1c of 7.9 implies that participants had relatively good control before entering one of the respective studies, which may suggest a largely adherent sample.

Although qualitative research is not considered to be necessarily generalizable by some audiences [47], it is incumbent on the researcher to provide full demographic descriptions so that the generic reader from an interdisciplinary audience can decide about the transferability of findings to his or her own practice, research, or development context. Furthermore, trustworthiness of the findings can be clarified based on participants’ checking of data and peer review of data analysis. This was a shortcoming of some of the literature that was reviewed.

Implications for Health Care Practice

Engaged participants spend considerable time on diabetes care, so it is important that they receive support to make informed choices. On the basis of this review, it was found that these are the people most likely to benefit from the affordances of advanced technology; however, this creates a tension between these populations and hard-to-reach groups who may be at increased risk of diabetes complications. Furthermore, Lawton et al [44] suggest that in general HCPs lack knowledge about the scope and purpose of advanced technology for diabetes. This is important, given the potential information overload and the frustrations that adults with T1D are presented with when using technology demonstrated within this review and other literature [9,19,41].

What is meaningful for the adult with T1D might not be important for the clinician and may therefore require mediation. Storni [43] found that patient-generated tags for mobile devices developed by participants were not of interest to clinicians who were more focused on numerical values.

James et al [47] have explored the perceptions and experiences of diabetes educators when supporting the use of advanced technology and suggest that there are challenges for all parties. This includes device costs, access to Wi-Fi, and appropriate mobile devices. CSII puts demands on diabetes services, and there are also challenges associated with keeping up to date with technology, such as the skills to analyze data from patient mobile devices. This research study suggested that there is a need for mentorship of HCPs and a review of service configurations as technology advances.
Implications for Future Design of Technologies

Engaging people in the design of technology for T1D is essential for meeting the requirements of the user. Within this review, O’Kane et al [42] suggested that the design of devices needs to be both discrete and more public for context-dependent behavior. Lawton et al [44] suggested that voice recognition for entering data would make data logging practices easier for some people. Engaged participants appeared to be able to deal with the hassle and time required for diabetes-related tasks. However, a challenge for designers is to build in time-efficient capabilities.

Implications for Future Research

There is a requirement for studies within the context of day-to-day data logging that are representative of the general adult population with T1D. There is more scope for research that explores how technology could be used to engage hard-to-reach groups. Many of the studies in this review were short-term; however, the study undertaken by Lawton et al [44] on the use of insulin bolus calculators was in-depth and over time (1 year), thus providing a rich and diverse view of adherence and nonadherence along a trajectory, which provided important nuances about human factors. There is also a need to study the role of significant others within data logging processes [41]. There appeared to be a dearth of mixed-methods studies, which if conducted through a rigorous methodological process have the potential to capture the complexity of human factors by maximizing the advantages of more than one research design. There is also a need for future studies that explore the sociocultural and demographic factors associated with technology uptake.

Limitations

A limitation of this review is that the data were drawn from databases, which excludes emergent unpublished research in a fast-moving field. However, this was mitigated by extracting the data from sources retrieved from 9 key databases covering the fields of health, medicine, and computer science, and the search was performed in collaboration with a university health care librarian.

Comparison With Prior Work

The application of critical appraisal frameworks used in this review made it possible to evaluate the reliability, validity, and trustworthiness of each of the studies under consideration. This review presents a contribution to the field in comparison with descriptive mapping reviews and highlights areas where research design could be improved. By abstracting data from each of the studies, it was possible to compare the findings and focus on the human factor constructs of adult populations with T1D, including older people.

Conclusions

The purpose of this systematic review was to explore the relationship between human factors and the adherence to technology for data logging in adults with T1D. The research design was an integrative review, given the interdisciplinary nature of research in the field and the diverse methodological approaches taken to inquiries.

The aim of the review was to analyze the relationship between human factors and adherence to technology for data logging in adults with T1D. Overall, the sample was drawn from homogeneous populations that may not be the complete representation of adults with T1D. Inconclusive evidence was found about the relationship between adherence to data logging with advanced technology and measurable outcomes. There was some suggestion that adults with T1D may feel more satisfied with their treatments on transition to advanced technology. Qualitative research suggested that the way in which technology is used varies along a continuum and is contingent upon the sociocultural context in which technology is used. As technology continues to advance, there is a need for more research into how trusting the individual is of personal treatment advice, which is generated through advanced technology.

Acknowledgments

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement #689810. The authors wish to thank the members of the PEPPER group: project management team (D Brown, F Torrent-Fontbona, P Herrero Vinas, L Nita, J Masoud, J M Fernández-Real, and B López) and coinvestigators (A Aldea, M Fernández Balsells, P Gay, P Georgiou, B Innocenti, Y Leal, N Oliver, R Petite, M Reddy, C Roman, J Shapley, and M Wos). We also would like to thank Katherine Staples, subject librarian at Oxford Brookes University.

Authors’ Contributions

MW was the principal author of this review. The contribution included writing the protocol, undertaking the search strategy, first stage critical appraisal of all studies in the review, data analysis, abstraction and synthesis of the studies, and production of a first outline draft of the manuscript. CM contributed by overseeing and revising the section on advanced technologies and second peer review of all the critical appraisal of all the studies featured in the review. RF contributed by overseeing the results of the review and third peer review of the critical appraisal of all the studies featured in the review. DD contributed by overseeing and revising the section on human factors. RH contributed by peer-review of the manuscript and writing the conclusion to the review.

Conflicts of Interest

None declared.
Multimedia Appendix 1
Systematic review protocol.

[PDF File (Adobe PDF File), 49KB - humanfactors_v5i1e11_app1.pdf]

Multimedia Appendix 2
PRISMA flow diagram.

[PDF File (Adobe PDF File), 51KB - humanfactors_v5i1e11_app2.pdf]

Multimedia Appendix 3
Data abstraction, technology, and human factors.

[PDF File (Adobe PDF File), 87KB - humanfactors_v5i1e11_app3.pdf]

References


42. Storni C. Complexity in an uncertain and cosmopolitan world. Rethinking personal health technology in diabetes with the tag-it-yourself. PyschoNology J 2011 Sep 01;9(2):165-185 [FREE Full text]


Abbreviations

- BG: blood glucose
- CGM: continuous glucose monitoring
- CSII: continuous subcutaneous insulin infusion
- CLC: closed loop control
- HbA1c: glycated hemoglobin
- HCP: health care professional
- MDI: multiple daily injections
- PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
- SAPT: sensor-augmented pump therapy
- T1D: type 1 diabetes

©Marion Waite, Clare Martin, Rachel Franklin, David Duce, Rachel Harrison. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 15.03.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
Reducing Misses and Near Misses Related to Multitasking on the Electronic Health Record: Observational Study and Qualitative Analysis

Neda Ratanawongsa1,2, MPH, MD; George Y Matta1,2, BA; Fuad B Bohsali3, MD; Margaret S Chisolm4, MD

1Division of General Internal Medicine, Department of Medicine, University of California, San Francisco, San Francisco, CA, United States
2University of California, San Francisco Center for Vulnerable Populations at Zuckerberg San Francisco General Hospital, San Francisco, CA, United States
3Division of Hospital Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States
4Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Abstract

Background: Clinicians’ use of electronic health record (EHR) systems while multitasking may increase the risk of making errors, but silent EHR system use may lower patient satisfaction. Delaying EHR system use until after patient visits may increase clinicians’ EHR workload, stress, and burnout.

Objective: We aimed to describe the perspectives of clinicians, educators, administrators, and researchers about misses and near misses that they felt were related to clinician multitasking while using EHR systems.

Methods: This observational study was a thematic analysis of perspectives elicited from 63 continuing medical education (CME) participants during 2 workshops and 1 interactive lecture about challenges and strategies for relationship-centered communication during clinician EHR system use. The workshop elicited reflection about memorable times when multitasking EHR use was associated with “misses” (errors that were not caught at the time) or “near misses” (mistakes that were caught before leading to errors). We conducted qualitative analysis using an editing analysis style to identify codes and then select representative themes and quotes.

Results: All workshop participants shared stories of misses or near misses in EHR system ordering and documentation or patient-clinician communication, wondering about “misses we don’t even know about.” Risk factors included the computer’s position, EHR system usability, note content and style, information overload, problematic workflows, systems issues, and provider and patient communication behaviors and expectations. Strategies to reduce multitasking EHR system misses included clinician transparency when needing silent EHR system use (eg, for prescribing), narrating EHR system use, patient activation during EHR system use, adapting visit organization and workflow, improving EHR system design, and improving team support and systems.

Conclusions: CME participants shared numerous stories of errors and near misses in EHR tasks and communication that they felt related to EHR multitasking. However, they brainstormed diverse strategies for using EHR systems safely while preserving patient relationships.

(JMIR Hum Factors 2018;5(1):e4) doi:10.2196/humanfactors.9371

KEYWORDS

electronic health records; physician-patient relations; patient safety
Introduction

Clinicians spend one-third of outpatient visits using electronic health record (EHR) systems, either in silence or by multitasking [1-3]. Multitasking occurs when someone performs 2 or more tasks simultaneously. Common examples of clinician EHR multitasking are eliciting a history while entering data (voluntary multitasking) and listening to a patient’s question that arises while the clinician orders a prescription (externally prompted multitasking) [2,4]. Multitasking may increase the risk of making errors, either in communication with patients or in completing EHR tasks, such as documentation or computerized order entry [4-7]. Meanwhile, using EHR systems in silence has been associated with lower patient satisfaction [1,2]. However, delaying EHR system use until after visits may increase clinicians’ EHR workload, stress, and burnout [8,9]. This study describes the perspectives of clinicians, educators, administrators, and researchers about their experiences with misses and near misses that they felt were due to clinician multitasking while using EHR systems.

Methods

This observational study was a thematic analysis of perspectives elicited during 3 continuing medical education (CME) courses in 2017. Participants included clinicians, clinician-educators and -administrators, and researchers attending 90-minute workshops at international health communications conferences (23 participants in Rhode Island and Maryland, USA), and clinicians and allied health professionals attending a 45-minute lecture during a course on caring for vulnerable populations (40 participants in California, USA). Workshops began with storytelling exercises about memorable times when multitasking EHR use was associated with “misses” (errors that were not caught at the time) or “near misses” (mistakes that were caught before leading to errors). Workshops and the lecture included a literature review about multitasking [1,3], video reenactments from a recent study [3], and a visioning exercise about reducing multitasking errors.

One workshop facilitator (NR) transcribed notes and quotes from participants during the interactive portions of the sessions. Two researchers (NR, MSC, or GYM) used an editing analysis style to identify “meaningful units or segments of text that both stand on their own and relate to the purpose of the study” [10]. In these data, individual quotes could represent more than one concept and be categorized by researchers under multiple different codes. We came to consensus in codes and themes and then selected representative quotes. Between the second and third session, no unique codes or themes arose, and we deemed we had reached theoretical saturation [10].

A University of California, San Francisco (UCSF) Committee on Human Research granted an exemption for this evaluation.

Results

All workshop participants shared stories of misses or near misses (Table 1) in EHR system ordering and documentation or patient-clinician communication, wondering about “misses we don’t even know about.” Table 1 shows risk factors emerging from these stories.

Participants wanted strategies for using EHR systems during visits, while ensuring patients feel respected and heard. One participant lamented that “I’m torn between real and ideal. We would spend all day finishing notes, but [pretending to type while speaking] ‘Three sexual partners?’” Another shared that “If it has emotional value, they won’t tell me while I’m typing.” Strategies to reduce multitasking EHR misses included (Table 2) clinician transparency when needing silent EHR use (eg, for prescribing), narrating EHR system use, patient activation during EHR system use, adapting visit organization and workflow, improving EHR system design, and improving team support and systems.

When asked for take-home intentions, 1 clinician wished to be authentic in voicing his desire to “be on the same side” with patients, acknowledging the need to use the EHR system but saying “I don’t want it to get in the way. I want you to always be able to call me back to the present.”
<table>
<thead>
<tr>
<th>Themes and codes</th>
<th>Examples or quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Types of misses and near misses</strong></td>
<td></td>
</tr>
</tbody>
</table>
| EHR errors in ordering and documentation | - Prescription electronically sent to the wrong pharmacy: “Especially when I’m calculating pediatric doses. I do it right then and don’t want to make a mistake.”  
- Copied and pasted charting in the wrong chart: “Before you couldn’t easily get into someone else’s chart accidentally...because you would have to pull the chart and open it.” |
| Communication errors | - “My agenda has changed unconsciously from my agenda or my patient’s agenda before to an agenda hidden to me that affects my focus, causes me to miss things in general.”  
- Unseen misses: “Record gives us a false sense of security that we’re capturing so much.”  
- “I suspect I am missing things, but I hope I’m catching the ‘red flags.’” |
| **Risk factors** | |
| Computer position | - “I’m worried I don’t even know that I’m missing something because my back is to the patient.” |
| EHR usability | - “The buttons are so close together that I can easily click into the wrong place.”  
- Time lags or glitches in the EHR program  
- “I now have to find the correct lab in computer and link to a diagnosis...nothing can go forward...”  
- “If you’re not proficient in using the computer, it’s just hard and takes more time.” |
| Note content or style | - “[EHR] was set up to bill, but not really designed for communicating what’s important for patient care.”  
- “Before I could just draw a line down the pediatric physical examination boxes. Now I have to check each of multiple boxes.”  
- Information that is not useful: “dates when medications are filled” or “inaccurate problem lists”  
- Agenda driven by EHR: “Conversation is being driven by something else.” |
| Information overload | - “Prerounding helps, but there’s just so much information now.”  
- “I find I’m going down more ‘rabbit holes’ for more information.”  
- More graphs and tools to use |
| Workflow | - Keeping multiple patient charts open simultaneously  
- Interruptions by other team members  
- Inability to delegate: “I used to be able to ask someone to help me, but I have to do it myself now.” |
| Systems issues | - High volume and short visits: “I can’t imagine what the surgical specialties must do.”  
- Perceived urgency for documentation: “Pressure for immediacy...it’s an unacceptable delay.”  
- Concern about adding to EHR use after hours: “It will be 3 hours of my life later.” |
| Provider and patient communication behaviors and expectations | - “Monologue style of communication” without “open-ended invitations”  
- Verbal “uh-huh...trumped by nonverbal body language” suggesting provider not listening  
- Patients interrupting silent EHR use: “They think they can talk and that you can hear and listen to them, but you can’t.”  
- Culture of screens: “It’s normal to have your face in a screen...maybe more typical more so than normal.” |
### Table 2. Strategies elicited from continuing medical education conference participants for reducing misses and near misses due to multitasking on electronic health records (EHRs).

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Examples or quotes</th>
</tr>
</thead>
</table>
| Awareness and transparency when silent EHR use needed | • “Previewing is always helpful. There are times today when we’re going to be talking 1:1, and there are times when I’ll be using this computer. Sometimes I may even have to use the computer quietly, and while I’m doing that, you can be doing this.”
• “Like in the hospital, where some nurses wear a ‘stop’ sign vest for med pass—they worried about patients minding it, but when they explain it as a ‘safety measure’ then patients understand.” |
| Narrating during EHR use                | • “I talk out loud when I’m looking up test results, and I interpret the results for them. I think it helps to know what I am doing and educates them, too.”                                                                 |
| Patient activation during EHR use      | • Invite History of Present Illness/Review of Systems completion together: “check these boxes with me.”
• Give patient education handout to review
• “While I’m putting this in the computer, why don’t you write down what we talked about [or] what you’re going to work on before the next visit.”
• “How will you remember this? Why don’t you think about that and we’ll talk afterwards.”
• Invite patient to “Call me back to the present.” |
| Visit organization and workflow         | • Preround before visit
• Avoid using the computer at the beginning of the visit or during sensitive conversations
• “I’m going to try to bunch things together to avoid going in and out and back in to the same section again. Like trying to do all the meds at the same time.”
• Ask patients to prepare for examination (eg, removing footwear for diabetes foot examination or undressing child for pediatric well visit) |
| Improving EHR design                   | • Make displays of patient photos accessible for safety to reduce wrong chart documentation
• Reduce structured data to allow narrative documentation |
| Team support and systems                | • Voice recognition documentation support
• Medical scribe support: “When I saw my doctor the last time, she had a resident typing for her, and it was like a different world. She was actually looking at me.”
• Team support in visit documentation: “If I had help, I’d much rather have med rec before and help linking labs to ridiculous diagnoses…”
• “We’re being measured on patient satisfaction and quality outcomes. Both are being measured, and so both of those may be more important than doing administrative work.” |

### Discussion

CME participants shared numerous stories of errors and near misses in EHR system tasks and communication that they felt related to EHR multitasking. However, they also brainstormed diverse strategies for using EHR systems safely while preserving patient relationships.

Clinicians need practical intrapersonal, interpersonal, and systems strategies to use EHR systems in mindful, relational ways. Avoiding all EHR use during patient encounters may be impossible and unsustainable, with clinicians using EHR systems over half of their workday and increasingly after clinic hours [8,9]. Meanwhile, research suggests that the risk of EHR multitasking is affected by the cognitive complexity of tasks and decisions, EHR system usability, teamwork, and clinician-patient dynamics [2-7].

Clinical multitasking predated EHR systems, which can reduce the risk of making errors by reducing the cognitive load of clinicians’ work by synthesizing and organizing information in accessible, usable formats. A 2009 Israeli study found that clinicians perceived some benefits to reducing the cognitive load of completing some clinical tasks, particularly if they perceived the EHR system to be comprehensive and usable [5].

At the same time, a danger of growing comfort and automaticity with EHR use was a risk of medication or documentation error [5]. More recent research has suggested that medication errors and adverse drug events in intensive care, hospital, and ambulatory settings may be reduced with computerized provider order entry and drug-drug interaction checking [11-13], although
continuing research about errors and near misses with computerized provider order entry may yield further improvements to reduce the cognitive complexity of EHR ordering [14]. This study adds to this growing literature in the context of the rapid expansion of newer-generation EHR systems in the United States under the meaningful use incentives programs.

Clinician transparency with patients about using EHR systems—including tasks such as prescribing that require focused attention to avoid errors—may result in fewer misses while preserving patient trust and satisfaction. As professional schools implement skills-based training in patient-provider communication with EHR system use [15], trainees may be able to practice empathic ways to negotiate the need for silent EHR use and ways to detect subtle queues from patients signaling that they need the clinician’s full attention.

In addition, other systematic approaches are needed to mitigate technology-induced errors—that is, medical errors arising from a technology’s design and development, implementation and customization, and resultant human-computer interactions and sociotechnical work processes [6,7]. These include slips (errors that are corrected) and mistakes (errors that go unnoticed or uncorrected) [7]. Borycki recommended proactive and reactive methods for reducing technology-induced errors: heuristic evaluation, cognitive walkthroughs, usability testing, clinical and computer-based simulations, rapid assessment processes, ethnographies, and case studies [7].

Future studies should explore diverse patient perspectives about clinicians’ EHR multitasking and their strategies for bringing clinicians “back to the present.” In addition, studies should examine how these strategies affect patient-important outcomes in quality and safety.

Acknowledgments
This research was supported by US Agency for Healthcare Research and Quality (AHRQ) grants 1K08HS022561 and P30HS023558. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of AHRQ.

Authors’ Contributions
NR had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. NR, GYM, and MSC conceived of the study concept and design. NR, GYM, FBB, and MSC acquired the data. NR, GYM, and MSC drafted the manuscript. NR, GYM, FBB, and MSC interpreted the data. NR, GYM, and MSC conceived of the study concept and design. NR, GYM, FBB, and MSC acquired the data. NR, GYM, and MSC critically revised the manuscript. NR, GYM, and MSC critically revised the manuscript. NR obtained funding. GYM provided administrative, technical, or material support. NR supervised the study.

Conflicts of Interest
None declared.

References


Abbreviations

CME: continuing medical education
EHR: electronic health record

©Neda Ratanawongsa, George Y Matta, Fuad B Bohsali, Margaret S Chisolm. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 06.02.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.

Ned T Sahin1,2, PhD; Neha U Keshav1, PhD; Joseph P Salisbury1, PhD; Arshya Vahabzadeh1,3, MD

1Brain Power, Cambridge, MA, United States
2Department of Psychology, Harvard University, Cambridge, MA, United States
3Massachusetts General Hospital, Boston, MA, United States

Corresponding Author:
Ned T Sahin, PhD
Brain Power
1 Broadway
Cambridge, MA,
United States
Phone: 1 617 514 7300
Email: sahin@post.harvard.edu

Abstract

Background: Computerized smartglasses are being developed as an assistive technology for daily activities in children and adults with autism spectrum disorder (ASD). While smartglasses may be able to help with educational and behavioral needs, their usability and acceptability in children with ASD is largely unknown. There have been reports of negative social perceptions surrounding smartglasses use in mainstream populations, a concern given that assistive technologies may already carry their own stigma. Children with ASD may also have a range of additional behavioral, developmental, and social challenges when asked to use this emerging technology in school and home settings.

Objective: The usability and acceptability of Glass Enterprise Edition (Glass), the successor to Google Glass smartglasses, were explored in children with ASD and their caregivers.

Methods: Eight children with ASD and their caregivers were recruited to attend a demonstration session with Glass smartglasses the week they were publicly released. The children had a wide range of ability, including limited speech to speaking, and represented a full range of school ages (6 to 17 years). Children and caregivers were interviewed about their experience of using the smartglasses and whether they would use them at school and home.

Results: All 8 children succeeded in using Glass and did not feel stressed (8/8, 100%) or experience any overwhelming sensory or emotional issues during the session (8/8, 100%). All 8 children (8/8, 100%) endorsed that they would be willing to wear and use the device in both home and school settings. Caregivers felt the experience was fun for the children (8/8, 100%), and most caregivers felt the experience was better than they had expected (6/8, 75%).

Conclusions: A wide age and ability range of children with ASD used Glass immediately after it was released and found it to be usable and acceptable. Despite concerns about potential stigma or social acceptability, all of the children were prepared to use the technology in both home and school settings. Encouragingly, most caregivers noted a very positive response. There were no behavioral, developmental, or social- or stigma-related concerns during or after the session. Smartglasses may be a useful future technology for children with ASD and are readily accepted for use by children with ASD and their caregivers.

(JMIR Hum Factors 2018;5(1):e1) doi:10.2196/humanfactors.8785

KEYWORDS

autism; technology; digital health; augmented reality; virtual reality; smartglasses; usability; schools; education; classroom; IDEA; IEP; special education
Introduction

Background
Autism spectrum disorder (ASD) is a childhood-onset developmental disorder, with an estimated 3.5 million people being diagnosable with ASD in the United States alone [1]. Innovative assistive technologies may help to address the unmet educational and therapeutic resource demands of the ASD community [2]. While there are many different types of assistive technology, the portability, capability, and ubiquity of smartphone and tablet devices has led to considerable growth in assistive apps for these devices [3,4]. More recent technological advances have resulted in lightweight smartglasses: face-worn computers with a visual display and in-built sensors [5-7] that can also deliver assistive apps [8,9].

Smartglasses can deliver a large range of experiences, including augmented and virtual reality [10]. They are also sensor-rich and can collect a wide range of quantitative user data [9,11,12]. These data can be monitored and analyzed on a real-time basis, allowing for the smartglasses to dynamically change the user experience to optimize learning—effectively placing the user and the smartglasses in a closed feedback loop [8,13,14]. Given the proximity of smartglasses to the sensory organs contained in the human head, this type of computing may enable a higher level of human-computer interaction than other devices [13]. Smartglasses are already being developed as a social and behavioral communication aid for people with ASD [8,15,16].

There are a number of important differentiating factors to consider when smartglasses are compared to handheld devices. Handheld devices such as tablets and smartphones require one or both hands to hold the device and encourage a heads-down posture (Figure 1 A, left) [17]. Evidence suggests that smartphone use may decrease user awareness of their social and physical environment. This is a particular concern in people with ASD, given that they already often face challenges engaging with the social world around them [18]. In contrast, head-worn computers pose an advantage in allowing and potentially encouraging children to remain heads-up while using them. This heads-up posture when using smartglasses can allow for better user engagement with people and the social world (Figure 1 A, right).

Modern Assistive-Reality Smartglasses
The emergence of a new crop of smartglasses is encouraging, especially because the initial public reaction to the widely recognized original Google Glass resulted in some negative social reactions. Modern smartglasses vary in terms of physical dimensions, functionality, and intended user group. For the purposes of this report, we decided to investigate the acceptability and usability of the most recently released lightweight smartglasses, Glass Enterprise Edition (Glass). Glass was released by X (a subsidiary of Alphabet Inc, formerly known as Google X) in July 2017. Glass is an assistive-reality technology, and it is the successor to Google Glass, one of the most recognizable smartglasses in the world [19]. Glass, like its predecessor, is a head-mounted, wearable computer that has demonstrated utility in a variety of situations where operating a computer hands-free and while heads-up is of particular advantage. Glass has been creatively developed as a technology that can deliver social and cognitive skills coaching to children and adults with ASD [8]. To our knowledge, we have reported on the first studies of ASD-related software on the original Google Glass (Explorer Edition) [8,9,15,16], and here we present the first appearance of Glass (Enterprise Edition) in the literature.

It would seem that the Enterprise Edition (which has updates to the form factor, usability, central processor, display, audio system, and other features) would represent a substantial advantage for assistive technology apps and algorithms for ASD. However, it remains unknown whether people with ASD would actually desire to wear the new device. Assistive apps for people with ASD on the original Google Glass have been shown to be tolerable [20], safe [15], and to reduce hyperactivity in an ASD sample [8,16]. However, small changes in devices can greatly affect the desire of potential users to wear them.
Given that the initial entry of Google Glass and other smartglasses raised privacy concerns and some negative public reaction, the announcement of a major new release of head-worn computing [19] signaled a potentially major advance for assistive technology targeting populations who traditionally face significant social challenges [17]. Google Glass was ahead of its time and may have been held back by perceptions around desirability and social acceptability of wearing this new category of device in public [21,22]. It is therefore reassuring to developers that head-worn computer platforms have received public backing from one of the largest companies in the world [19], in this case the inventor of the product [23].

Understanding the Needs of People With Autism Spectrum Disorder

As with any assistive technology, it is important to investigate and understand the attitudes of children and young adults with ASD, especially because children with special needs are often forced to use devices and systems they do not actually like or want to be associated with [24,25]. This is ultimately less effective because aversion leads to lower adherence. Poor adherence and problems with maintaining lasting engagement are some of the largest issues facing educational devices and apps as well as well-being and lifestyle tools [26,27].
Many people with ASD use assistive technology to help them with communication skills, social and emotional skills, and adaptive/daily activities and living skills [28]. Assistive technologies elicit a range of responses from individuals and their peers, and they can be considered cool [25], weird, desirable, or a source of stigma [29,30]. Users of assistive technologies can often express a preference for the type of assistive technology that they want to use [31,32], even at a young age [33]. Additionally, the social acceptability of an assistive technology may be one of the most important elements in determining if that technology gets used by people with developmental disabilities [30,34]. These individuals have often had to use technologies that have been selected for them and their families while having little input to the potential negative image, stigma, or embarrassment of using such technologies [30]. Understanding and implementing user preference of assistive technologies empowers self-determination in these individuals [31]. The preferences and views of the family and caregivers of these individuals are also important as they impact the acceptance and effective use of such technologies in the household [28,35]. These issues are pertinent to smartglasses in light of past reports of negative public perception (eg, around privacy concerns [22]).

There have only been a handful of reports on the use of smartglasses in people with ASD [8,15,16], and the attitudes toward and acceptability of such devices to people with ASD remains unclear. The use of smartglasses in people with ASD also requires discussion of their potential impact on social communication from a cognitive neuroscience standpoint and their prospective influence on child development from ecological, psychosocial, and cognitive child development theories.

**Potential Impact of Smartglasses on Social Communication**

The human face, a complex and dynamic system, is our most powerful means of social communication [36]. To successfully transmit social information to another person, the sender must have the mental and physical means of generating a facial and bodily representation of the social information that she or he wishes to send, while the receiver must be in a position to see and decode the facial and bodily representations into social information. The social communication deficits seen in ASD may impede the ability to send and receive social information. People with ASD are reported to have deficits in facial perception [37,38], emotion recognition [39], eye gaze [40], and production of facial expressions [41]. It is important to consider the possibility that social communication may be further impacted by the physical presence of smartglasses on a sender’s face. Smartglasses may impede social communication if, for example, the sender demonstrates a hesitancy in producing natural head movements or expressing large magnitude facial emotional expressions due to concern that the smartglasses may fall off the face or be damaged. Smartglasses may also impair social interaction if the user feels the assistive device is socially undesirable [42] or a source of stigma [30]. In these situations, users may not use the device or may alter their facial and bodily actions to minimize attention to themselves. Furthermore, the physical form factor of smartglasses may obscure a portion of the wearer’s face that is visible to others, especially the central information-rich parts of the face such as the eye regions [43].

The relative effect of this obscuring of the facial region may be dependent on the size of the individual’s face relative to the smartglasses, which may correlate with the age of the individual given that biologic age determines an individual’s head size [44]. It may also depend on the ability of the receiver to successfully compensate for partly missing facial data and to make inferences about a sender (a common application of this in ASD research is the “Reading the Eyes in the Mind” test [45]). Since people with and without ASD find it more difficult to read the facial emotional expressions of people with ASD [41], it is conceivable that further obscuring the amount of visible facial information could make the interaction even more arduous. This point may be particularly relevant to interactions between people with ASD and their unaffected family members. ASD is a highly hereditable condition with a complex genetic basis [46], and many unaffected relatives of children with ASD have been found to have subclinical autistic traits [47]. The parents of children with ASD may demonstrate subtle deficits in social communication and face processing [48,49].

Given these reports and considerations, the physical presence of smartglasses may affect social communication, and it may be sensible to attempt to minimize such facial obscuration to enhance social communication between people with ASD and their family members.

The presence of face-worn smartglasses may also influence social relationships, of the adults or children who wear them, as they alter a user’s facial appearance. Unlike many other assistive technologies, they are not easy to hide. Wearing smartglasses may not only alter how the user perceives the world but may alter how the world perceives the user. Facial appearance plays a key role in determining how people interact with one another [50], including whom they help, hire, or want to date [51]. Human faces may also be judged based on their symmetry, a marker of attractiveness and an indicator of optimal developmental outcome despite environmental stressors [52]. Greater facial symmetry has been linked to increased perceived trustworthiness and a decreased risk of being bullied [53]. Facial symmetry may be perceived as demonstrating genetic quality and therefore suitability of an individual as a mate [52], while facial asymmetry may be a predictor of long-term psychological, emotional, and physiological distress [54]. Users of smartglasses that are asymmetrical, such as those that are monocular, could be perceived as being less attractive and trustworthy due to the aforementioned principle of evolutionary psychology. By extension, “asymmetric” smartglasses users may also be at greater risk of bullying [53]. On the other hand, smartglasses that are asymmetrical may obscure less of the wearer’s face from the view of others. As discussed earlier, maximizing how much of the face is visible may help facilitate social communication. Even nontechnological face-worn glasses are associated with impaired interpersonal relationships: for example, wearing prescription glasses or having a history of using eye patches has been associated with a 35% increase in the likelihood of receiving physical or verbal bullying [55].
Smartglasses in the Context of Child Development

The perceptual impact of smartglasses and their ability to augment a child’s cognitive and emotional functioning may have a central and influential role in childhood development if we consider Bronfenbrenner and Ceci’s bioecological model [56] and Bronfenbrenner’s earlier ecological systems theory [57]. According to the bioecological model, children are active participants in their environments and they have unique bidirectional interactions with each of their contextually separate environments, including home and school. This model places increased emphasis on the cognitive, emotional, and physical attributes of the child in his or her development and in how the child and environments interact with one another. As outlined in Bronfenbrenner’s ecological systems theory [57], the school environment, like the home environment, is one of the most intimate and influential environments affecting childhood development, as it lies in the child’s microsystem. When we consider that smartglasses may enhance the cognitive and emotional functioning of children within their microsystem, we can see that they may have a highly influential role in child development. Even within the microsystem, the contextual differences between the most intimate environments may affect a child’s view toward using assistive technology. Research has shown that children have different attitudes and levels of enthusiasm toward using assistive technology depending on whether they are asked to use it at home or at school [32].

Furthermore, use of smartglasses by future school-age children and adolescents should prompt a discussion of Erikson’s 4th and 5th psychosocial stages [58]. Erikson identified a range of psychosocial developmental stages from birth through death. School-age children experience Erikson’s 4th psychosocial stage, described as a psychosocial crisis of industry versus inferiority. A child in this stage is often expected to learn and demonstrate new skills, productively complete tasks, and meet the expectations of parents and teachers. During this stage, a child becomes aware of his or her abilities and the abilities of his or her peers. A child who cannot master these expected skills risks a sense of inferiority and failure. The potential impact of smartglasses on this developmental stage is not known. They may aid children in successfully mastering this psychosocial stage by allowing them to be productive and giving them a sense of achievement. There is also a risk that children may feel inferior if they feel that without the smartglasses they are incompetent or if they feel ridiculed for wearing such devices. Each child may face a unique situation based on his or her personal attributes and the support received from key people such as teachers, parents, and peers. This highlights the importance of ensuring that these key people are familiar with smartglasses technology and understand its capabilities and functionality.

Following this stage is Erikson’s 5th psychosocial stage that occurs in adolescence, described as a psychosocial crisis between identity versus role confusion [58]. Adolescence is a time of tremendous biological and psychological change [59], and during this stage individuals seek to define their role in the world, seeking to address the existential question, who am I and what can I be? Individuals will try to find like-minded social groups, focus on relationships with peers, and pursue a sense of belonging. Many questions remain unanswered about how smartglasses may impact people with ASD during this stage, especially given the many social challenges people with ASD encounter during this transition from childhood to adulthood [60]. How will ASD and these technologies define the individual? Will these technologies help individuals to find their purpose or hinder them? The impact of such technology may depend on smartglasses’ physical attributes, their impact on social relationships, or individual person characteristics (as discussed above within the scope of the bioecological model [56]).

Learning happens continuously in childhood, and the use of smartglasses technology may provide a digital means of enabling learning to occur, as in Vygotsky’s zone of proximal development (ZPD) [61]. Vygotsky originally described his ZPD as being “the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers” [62]. These smartglasses designed as assistive technologies may allow children to undertake and learn tasks that they would have found impossible or very difficult to do independently. A child with ASD normally has a number of challenges in being in the ZPD, such as becoming overwhelmed with new experiences, struggling with transitions in environment or activities, and coping with sensory stimuli [18]. Sensor-rich smartglasses may be of particular utility here in that they can be used, with the right software, to monitor the behavioral and physiologic functioning of a child. For instance, they can be transformed by software to be able to detect when children are under- or overstimulated and to accordingly adapt the learning experience in real time to keep a child engaged and in the ZPD [8].

Victimization, Socialization, and the School Environment

School-age children with ASD are at risk of being stigmatized [63] and being victims of bullying [64] for multiple reasons. They have different developmental trajectories that may put them at greater risk of victimization than their neurotypically developing peers, especially when they have challenges in social skills and communication [64]. They may struggle to recognize social cues and develop relationships with their peers, impeding their ability to be better integrated by the community [65-67]. Bullying may be particularly problematic at school, where academic and social factors may be a source of considerable stress, anxiety, and mental health concerns in children [68-70]. A school represents not only an academic establishment but a complicated and highly social environment. Children in schools often balance interpersonal relationships with peers and staff, complex social hierarchies, and school rules that can dictate the most basic elements of children’s day (whom to play with, where to sit, and when to talk to others [65-67,71]). Some reports have suggested that children with ASD have inherently low motivation or desire to join social groups, but recent evidence indicates this is not the case and many have a strong desire for acceptance [72-74]. Therefore, it is important to consider the acceptability and design of any assistive device in...
the population, given the risk of stigma and social isolation [30]. This is especially true for a device that is worn on the face.

Methods

Study Outline

We gave 8 children with ASD an opportunity to try the Glass smartglasses in a controlled, recorded environment and to explore its features, usability, and visual characteristics. We observed and recorded the interaction of the children with the device. We also conducted a postsession semistructured interview with the children and their caregivers, who accompanied the child and observed the whole session. Our sample represented a broad age range and severity spectrum of ASD.

Institutional Review Board Statement

The use of the Brain Power Autism System running on multiple head-worn computing devices by children and adults with autism was approved by Asentral Inc Institutional Review Board, an affiliate of the Commonwealth of Massachusetts Department of Public Health. The study was performed in accordance with relevant guidelines and regulations.

Participants

Eight children with clinically diagnosed ASD and their caregivers were entered into this study. The participants represented a wide range of school-aged children, ages 6.7 to 17.2 years (mean 11.7 [SD 3.3] years), including 7 males and 1 female. Participants were recruited from a user research database created from Web-based research interest forms. Written consent for study participation was obtained from the legal guardians, and children from age 7 to 17 years provided written assent. In this report, every participant was accompanied by a parent or guardian caregiver to the session, and participants and caregivers could exit the session at any time and for any reason. It was explained that the main aim of the study was to understand the acceptability and usability of modern smartglasses technology in children with ASD.

Caregivers rated the participant level of overall ASD functioning according to a subjective 7-point scale (1=lowest-functioning/severe to 7=highest-functioning/mild). Caregivers also rated speaking ability on a similar scale (1=nonspeaking to 7=fully conversational). Participants represented a large range of overall ASD functioning (range 4 to 7 out of 7; mean 5.6 [SD 1.1]) and speaking ability (range 4 to 7 out of 7; mean 5.5 [SD 1.3]).

Data Collection Procedure

Participants and their caregivers were orientated to the testing room where they had an opportunity to learn about the Glass smartglasses and to physically wear and use them. They were provided with any assistance they required to properly place the smartglasses on their heads and align it with their eyes, although little assistance was needed. They were able to use any of the apps on the smartglasses. Testing sessions were recorded via video and photographs. All participants and/or caregivers gave written consent for their images and video to be used in current and future research analyses.

Following the testing, participants and caregivers went into a separate room where they were questioned about their experience as part of a semistructured interview. The participants were asked to compare their experience of Glass with previously tested assistive devices and gamified apps related to ASD. As previously noted, the participants were recruited from a research database for technology-related studies in ASD, and all had seen and tried the original Google Glass. Participants were asked if they became stressed when using the device and if the session was an overwhelming sensory or emotional experience for them. The questions were adapted or simplified based on the child’s speaking ability and were repeated if needed. Study staff interacted with the child and caregiver and spent time ensuring the questions were understood, considered, and accurately answered.

Participants were then asked whether they would consider wearing and using the device for 1 hour each day in their school and separately asked the same question about using the device at home. The caregiver was also interviewed in order to rate whether they felt the experience was fun for the participant and whether they felt the experience with the smartglasses went better than they had expected.

Exclusions

Individuals who had a known history of epilepsy or seizure disorder were not asked to take part in this study. Individuals who had any uncontrolled or severe medical or mental health condition that would make participation in the study predictably hazardous were also not invited to participate.

Results

All 8 children, who represented the full range of school ages (6 to 17 years), successfully wore, interacted with, and explored one or more Glass smartglasses (Figure 2). The smartglasses were loaded with a suite of assisted-reality apps for social-emotional learning and self-coaching related to brain-based challenges and needs, as discussed elsewhere [8]. Participants explored the devices at their leisure, putting them on and taking them off and exploring the style, size, weight, shape, and features such as foldability, and spoke out loud in some cases (children with greater speaking ability) about their observations and questions. All children successfully transitioned to the interview room, where they responded to questions by the experimenter, accompanied and assisted by their caregivers as needed. There were no negative effects reported or observed.

All participants noted that they did not feel stressed (8/8, 100%, Table 1) or have an overwhelming sensory or emotional experience when using the smartglasses (8/8, 100%). The participants all reported that they would be agreeable to using the smartglasses in both home (8/8, 100%) and school settings (8/8, 100%). Caregivers reported no concerns with the children using the smartglasses, and all caregivers reported that their child appeared to have fun using the device (8/8, 100%). The majority of caregivers felt the interaction of the child with the smartglasses went better than they had expected (6/8, 75%; Table 2). Of the remaining 2, 1 parent said that the experience...
had proceeded “as expected” and another answered the question conversationally but without a direct response, so the response was not tabulated as a yes but as an undetermined.

**Figure 2.** Children on the autism spectrum using and exploring the Glass Enterprise Edition device during a testing session at Brain Power. Each of the 8 participants, who represent the entire range of school ages, range from mild to moderate autism severity, and demonstrate a wide breadth of speaking ability (from moderate impairments in speech to being fully conversational), rated Glass Enterprise Edition as desirable to wear on their heads and use daily in the often-complex social environment of school and at home.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes n (%)</th>
<th>No n (%)</th>
<th>Neutral or undetermined response n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would you wear the smartglasses for 1 hour each day at school?</td>
<td>8 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Would you wear the smartglasses for 1 hour a day at home?</td>
<td>8 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Did you feel stressed while wearing the smartglasses?</td>
<td>0 (0)</td>
<td>8 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Did you feel overwhelmed (emotionally/sensory)?</td>
<td>0 (0)</td>
<td>8 (100)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>
Dissussion

Principal Findings

Smartglasses are an emerging technology that could hold much promise as an assistive technology for children and young adults with ASD. It is important to seek the opinions of children with ASD and their caregivers when considering the use of a new assistive device. This is especially true of smartglasses given their high level of visibility, prior reports of negative social perceptions, and the potential interplay of such devices with social communication and child development. Children with ASD and their caregivers may be particularly discerning about factors that could impact the use and social acceptance of such technologies in educational settings such as schools and in the home environment.

The results demonstrate that Glass was acceptable and desirable by all participants, who spanned the full range of school ages (6 to 17 years). It was encouraging to find that all 8 school-aged children with ASD felt that using these smartglasses was not a stressful experience and denied being overwhelmed in a sensory or emotional way. Additionally, it was also promising to see that all of the children expressed a willingness to use these devices in both school and home settings. Caregivers noted that children had fun using the device, and most caregivers felt their expectations of how the children would interact with the smartglasses were surpassed.

These results are important for a number of reasons. Children with ASD are frequently not involved in providing design or usability feedback to interventions and technologies developed for them. Involving children when choosing an assistive device is crucial to ensure that the device is socially appropriate for the environment, which will likely lead to greater compliance in wearing the device. It also appears that these children are accepting of new technologies, even on relatively uncommon and highly visible platforms such as head-mounted computers. The children who participated in this study were more open to using Glass in a public environment than many adults have been [22]. With this in mind, it will be equally as important to ensure caregivers and peers in the child’s microsystem are accepting of the assistive technology [57], as their opinions will likely sway a child’s enthusiasm toward the device. Many children in this study mentioned favoring Glass because of its unobtrusive, sleek design; having a device that is less noticeable and designed to be “cool” may help with its social acceptance and may not carry the stigma of assistive technology with it. The desirability of Glass in this case was predicated on a prediction of social acceptability (colloquially, the “cool factor”) in a social situation. Many factors may be included in a participant’s prediction of the cool factor of a device. Such factors may include unobtrusiveness, lightness, futuristic look, comfort, ease of storing, ease of transport, durability, ruggedness, styling, ability to give others experiences they could not otherwise have (conferring to the child an ability to control a social situation in a positive way), ability to initiate a conversation with decreased anxiety over selecting the topic of the conversation (ice-breaker), and more.

Limitations

The unanimous willingness of participants to wear the smartglasses in school is also important. The school setting is a place of high risk relative to social integration and stigma that could result from an undesirable or socially inappropriate device or behavior. This is one reason we chose the question of acceptability of the device at school as a high-bar test for how desirable and acceptable this new device may be. However, a limitation of this work is that we asked for the opinion of the target users, and such an opinion is necessarily based on a prediction. It may be hard to predict how a device or behavior will actually be received in the complex and changing social hierarchy of a school environment. Additionally, children with ASD may have extra challenges in predicting the emotional reactions and behaviors of their classmates, especially if they are in an integrated school environment with neurotypical or typically developing children their same chronological age. For all these reasons, further research is needed to test the acceptability within school environments.

Conclusions

These results suggest that a smartglasses platform may be an acceptable base for assistive software apps that could promote self-sufficiency. For instance, they may have a desirable new platform for gamified, social-emotional self-coaching apps based in neuroscience and artificial intelligence that have been deployed on other head-worn computer platforms [8]. The results are promising at a broader level for those who wish to use or develop apps that harness the unique features of this family of devices, such as their ability to allow the user to be heads-up, hands-free, and able to perceive and engage with the world around while receiving additional assistance. The results suggest that the newest entrant into the still-emerging family of devices may be well received, at least by some discerning populations. Further research is clearly needed to address these and more limitations or open questions of this work. This report represents part of a larger, ongoing research initiative.

This paper represents the first published work, to our knowledge, using Glass (Enterprise Edition). It also represents the first published use of Glass as an assistive or assessment device for people with different abilities or intellectual disabilities or challenges. This work extends our previous research on the use of the original Google Glass as an aid to people with ASD [8].

Table 2. Caregiver responses following use of smartglasses.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes n (%)</th>
<th>No n (%)</th>
<th>Neutral or undetermined response n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was it fun for your child to use Glass?</td>
<td>8 (100)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Did the experience go better than you anticipated?</td>
<td>6 (75)</td>
<td>0 (0)</td>
<td>2 (25)</td>
</tr>
</tbody>
</table>
Acknowledgments

The authors thank Google, Inc. for a generous grant and also the Glass team at X (formerly Google X) for technical guidance on Glass software development. This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Autism Research Program under Award No. W81XWH 17-1-0449. Early work to transform smartglasses into biomedical sensors was supported in part by the United States Army Medical Research and Materiel Command under Contract No. W81XWH-14-C-0007 (awarded to TIAx, LLC). Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the Department of Defense. The authors would also like to acknowledge Krystal Huey for her help with study coordination.

Conflicts of Interest

This report was supported by Brain Power, a neurotechnology company developing a range of artificially intelligent wearable technologies. Brain Power has engineering and technical partnerships with major technology companies and also receives funding support from federal and congressional sources.

Authors’ Contributions

NS is the inventor of the Brain Power Autism System. NS, JS, NK, and AV designed and undertook the intervention. The writing of this technology report was led by AV, and all authors contributed.

References


19. Kothari J. A new chapter for Glass. URL: https://blog.x.company/a-new-chapter-for-glass-c7875d40bf24?g=773a91e858b9 [accessed 2017-12-19] [WebCite Cache ID 6vpoUox7c]


53. Ocker M, Kreitler C. Facial symmetry perceptions in college students. CRIUS 2015;3(1).


Abbreviations

ASD: autism spectrum disorder
Glass: Glass Enterprise Edition
ZPD: zone of proximal development

Edited by G Eysenbach; submitted 18.08.17; peer-reviewed by S Shamsuddin, S Clough, I Montagni; comments to author 14.09.17; revised version received 08.10.17; accepted 29.10.17; published 04.01.18.

Please cite as:
Sahin NT, Keshav NU, Salisbury JP, Vahabzadeh A
JMIR Hum Factors 2018;5(1):e1
URL: http://humanfactors.jmir.org/2018/1/e1/
doi:10.2196/humanfactors.8785
PMID:29301738

©Ned T Sahin, Neha U Keshav, Joseph P Salisbury, Arshya Vahabzadeh. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 04.01.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
How Can Social Media Lead to Co-Production (Co-Delivery) of New Services for the Elderly Population? A Qualitative Study

Hadi Daneshvar\(^1\), MSc; Stuart Anderson\(^1\), BSc (Hons); Robin Williams\(^2\), PhD; Hajar Mozaffar\(^3\), PhD

\(^1\)School of Informatics, University of Edinburgh, Edinburgh, United Kingdom
\(^2\)Institute for the Study of Science, Technology and Innovation, University of Edinburgh, Edinburgh, United Kingdom
\(^3\)Business School, University of Edinburgh, Edinburgh, United Kingdom

Corresponding Author:
Hadi Daneshvar, MSc
School of Informatics
University of Edinburgh
IF-3.25, Informatics Forum,
10 Crichton St,
Edinburgh, EH8 9AB
United Kingdom
Phone: 44 7578561742
Fax: 44 131 6506513
Email: h.daneshvar@gmail.com

Abstract

Background: The future of health care services in the European Union faces the triple challenges of aging, fiscal restriction, and inclusion. Co-production offers ways to manage informal care resources to help them cater for the growing needs of elderly people. Social media (SM) is seen as a critical enabler for co-production.

Objective: The objective of this study was to investigate how SM—private Facebook groups, forums, Twitter, and blogging—acts as an enabler of co-production in health and care by facilitating its four underlying principles: equality, diversity, accessibility, and reciprocity.

Methods: We used normalization process theory as our theoretical framework to design this study. We conducted a qualitative study and collected data through 20 semistructured interviews and observation of the activities of 10 online groups and individuals. We then used thematic analysis and drew on principles of co-production (equality, diversity, accessibility, and reciprocity) as a deductive coding framework to analyze our findings.

Results: Our findings point to distinct patterns of feature use by different people involved in care of elderly people. This diversity makes possible the principles of co-production by offering equality among users, enabling diversity of use, making experiences accessible, and encouraging reciprocity in the sharing of knowledge and mutual support. We also identified that explication of common resources may lead to new forms of competition and conflicts. These conflicts require better management to enhance the coordination of the common pool of resources.

Conclusions: SM uses afford new forms of organizing and collective engagement between patients, carers, and professionals, which leads to change in health and care communication and coordination.

(JMIR Hum Factors 2018;5(1):e5) doi:10.2196/humanfactors.7856

KEYWORDS
social media; eHealth; mHealth; social networking; Web 2.0; health informatics

Introduction

Background

Health and care in the European Union faces the triple challenge of aging, fiscal restriction, and inclusion [1]. In the United Kingdom, the number of elderly people will increase to 6.6 million over the next 25 years. In Scotland, by 2035, the 65+ years age group is projected to have grown by 82%. This study focuses on three important problems related to the aging population in Scotland and the wider United Kingdom. The problems are (1) an increase in population of elderly people; (2) insufficient resources to meet the health and care needs of...
the elderly population; and (3) social exclusion of the elderly. These lead to an increased need for government expenditure to provide and deliver health and care services, as well as an increased need for expenditure by elderly people while their income is static or falling.

The statistics show the needs of elderly people are growing, and there is an increased requirement for carers [1]. Currently, the population of informal carers is more than 10% of the 65 million population of the United Kingdom. It is projected that this number will increase to 9 million of 73.2 million (around 12% of population) in the next 25 years. The current value of care is worth an estimated £132 billion per year—approximately equal to the total annual cost of health spending in the United Kingdom, which was £134.1 billion in year 2014-2015 [2]. So an important challenge is how to resource care and health of elderly people in the future. Depending solely on economic growth to fulfill the finance needs of public services is unlikely to meet the need in a time of austerity and will inevitably lead to poorer quality of service and outcomes. Hence new ways of meeting the need for health and care are needed [3]. To reshape service delivery, we need to consider how to utilize diverse resources.

The health and care system in the United Kingdom and Scotland is being reformed. The Scottish government has announced the need for better coordination and integration in this process [4]. Examining the concept of co-production is an initial step in reforming the service delivery. Boyle and Harris [3] from the New Economics Foundation give a definition for co-production:

Co-production means delivering public services in an equal and reciprocal relationship between professionals, people using services, their families and their neighbors. Where activities are co-produced in this way, both services and neighborhoods become far more effective agents of change.

There are a range of perspectives on the production and use of health and care services. A critical aspect of such services is the governance of their production and use. In this context, one strong standpoint sees health and care resources as “common pool resources” [5]. Common pool resources [5] refers to:

A system that is sufficiently large as to make it costly (but not impossible) to exclude potential beneficiaries from obtaining benefits from its use. To understand the processes of organizing and governing CPR [common pool resources], it is essential to distinguish between the resource system and the flow of resource units produced by the system, while still recognizing the dependence of the one on the other.

This common pool of resources may involve patients, informal carers, social carers, volunteers, professional carers (caregivers), and health professionals who can be seen as co-producers of health and care services. In this paper, we focus on informal carers, volunteers, and patients and examine how this large pool of informal carers and patients could, with more careful utilization, further augment the effort devoted to care in the United Kingdom. Current public services are poorly equipped to exploit the potential social economy of family and neighbors.

The full participation of informal carers in the co-production of health and care has the potential to play a significant role in the sustainability of health and care delivery. A pressing issue is how to coordinate this massive resource with the formal health and care system to enable true co-production of health and care. This massive resource is spatially dislocated and temporally uncoordinated and engaged in responding to very local demands. Modern information and communications technology (ICT) is viewed as a key enabler to overcoming such obstacles.

Increasingly eHealth and care services is viewed as the tool to reshape health care systems [6]. We propose that, in particular, social media (SM) can be viewed as an enabler for co-production. Communication is a key element in co-production that enables coordinating across various boundaries. SM cuts across boundaries, its use is well understood, but its effects are much more poorly understood. Therefore, this paper focuses on how SM enables this coordination.

To explore the role of SM in the context of co-production (with carers, patients, and volunteers in focus), we use Cahn’s framework as our analytical lens. Cahn [7] identifies the following principles as the elements that put co-production into action:

- Equality: no group or individual is more important than others. Everyone is equal and they have assets to contribute to the whole.
- Diversity: diversity and inclusion are important principles in co-production. So, diverse groups must be included.
- Accessibility: if everyone is going to take part on an equal basis, then everyone needs to have the same opportunity to be involved in activities, in a way that is suitable for them.
- Reciprocity: When people put in effort to contribute, they need to feel valued as well as needing to receive something back. This means that everyone is responsible and they have expectations, and therefore it is a mutual process.

Although these are critical elements of co-production, achieving all of them at the same time may result in asymmetry (among the elements) or conflicting goals. For instance, in some cases, encouraging inclusivity and diversity (having a large number and more diverse actors involved in one space) may be at the cost of equality and reciprocity (not everyone contributes equally or at all times). Therefore, in this paper, we initially highlight how SM enables these four elements, and then we discuss the possible conflicts.

By using this framework, we foreground the communication aspects of SM. We recognize this as one of the numerous aspects of co-production. In particular, further research is needed to explicitly heed to issues of resourcing, conflict and competition for resources, and the overall governance of health and care provision. Our focus here, therefore, is on the communication and cooperative utilization of health and care resources among patient representatives, carers, and volunteers. We will therefore discuss its limitations in the Discussion section and address the broader aspects and possible contentions involving health professionals and social workers in a later paper. SM are online tools for the creation and sharing of digital content. They aim
Kaplan and Haenlein [8] defined SM as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content.” Dissemination of content operates at Internet speeds. It has been argued that SM has caused a change in social action in many areas [9,10]. SM increases social interaction between patients and health organizations. Moorhead et al [11] explain that SM is a powerful tool for collaboration between users, and it acts as a social interaction mechanism for its wide range of professional and nonprofessional users. It empowers public and patients by enabling them to communicate with each other and exchange health information [9,10,12]. It enables users to discuss sensitive issues [13,14]. Moorhead et al conducted a systematic review of uses of SM for health and show that SM offers peer, social, and emotional support for its users. They also demonstrate that SM increases interactions for patients, their families, and their friends.

The term was coined by Shipley after his research and reports on technology trends [15]. SM has since become media for the creation and maintenance of connection and interaction among individuals [16,17]. They are currently used widely by a diverse range of users and have become among the largest number of most visited sites worldwide [18].

Several studies [8,19,20] categorize SM into 9 groups: (1) Wikis; (2) blogging; (3) microblogging; (4) content communities; (5) forum; (6) instant messenger; (7) social network sites; (8) mobile SM; and (9) virtual world and online social gaming.

In health and care, we divided SM to 3 groups: (1) SM that were created for general-purpose use and is now used for health and care, for example, Facebook groups (FBGs); (2) SM that were created for health purposes and make use of generic SM for other purposes (eg, 3D-Doctor or some other health applications (apps) that make use of Skype to connect patients to the doctors); and finally (3) apps created for health and care purposes that use the concept of SM (Health SM), for example, HealthUnlocked is new SM for health purposes [21].

**Aim**

The aim of our study was to investigate the current and possible future for SM as an enabler of co-production in health and care for elderly people. To achieve this aim, 2 main sets of questions are asked: (1) What are the uses of current SM in enabling (and reshaping) health and social care? and (2) How can SM be reshaped to enable (and reshape) health and care co-production?

We consider a typology of opportunities and limitations of SM for health and care. Figure 1 shows existing health and care service bundles with (1) existing or (2) new SM tools (new SM means some app developed for health and care that used the concept of SM such as HealthUnlocked). New health and care service bundles with (3) existing or (4) emerging SM tools. In this part of the research, we focused on (2): “How current SM help to reshape or change health and care services?” In other words, we assessed how existing SM acts as an intervention during the reshaping of health and social care in the United Kingdom by enabling co-production (in particular, co-services).

**Methods**

**Overall Project**

This paper focuses on one of the four aspects of a larger qualitative study that investigates the sociotechnical aspects of the current and possible future uses of SM by different organizations and groups of health and social care as an enabler of co-production in the United Kingdom, in particular Scotland. We conducted 20 semistructured interviews, which focused on the services offered, the types of online apps (particularly SM) used, their challenges, and the future possibilities of SM. We used purposeful sampling to select organizations and groups that were providing care services to elderly people. We used a combination of interviews and analysis of the activities and content of online groups to collect data. Combining different methods enabled us to triangulate the data sources to validate our findings.

**Material and Methods**

This paper focuses on the second section: How current SM help to reshape or change health and care services? (Figure 1). Our
appraisal adopts a sociotechnical technique [22,23], using a mixed-methods framework, including multiple methods (interviews, observations of online activities, and secondary data analysis) and multiple sources of data. Table 1 summarizes the data collection methods and sources. For the purposes of anonymity, names have been replaced with pseudonyms.

**Theoretical Framework**

Normalization process theory (NPT) has been used as our theoretical framework to enable us to obtain meaningful understanding of the complex sociotechnical processes involved in the use of SM tools and service within health and care co-production. NPT offers a whole system perspective, to assist researchers to make sense of the social and organizational aspects of different interventions and to better conceptualize the complex adaptive systems. NPT, which has been used in many eHealth research studies, has been used as a tool in this research to assess the changes brought about by the introduction of SM into the personal and organizational lives of patients, carers, and organizations involved in care activities (Table 2).

**Data Collection and Qualitative Data Analysis Procedure**

We conducted 20 interviews (approximately 22 hours) with patients, carers, and employees of third-sector or intermediary organizations (ie, charities), which provide funds and services for developing programs to reshape of health and care services in Scotland, and companies or organizations working in the health sector providing services to elderly patients with long-term conditions.

We used purposeful sampling to select the interviewees. Our purposeful sampling strategy aimed to identify organizations that actively used some type of SM in their activities. The selection criteria for organizational participants was people who were either involved in providing carer activities, decision makers, or those involved in design of ICT programs for elderly care. For nonorganizational participants, we aimed to select interviewees who were either patients or carers who actively used some type of SM in their day-to-day life.

We used the NPT framework to develop an open-ended interview question guide (Table 3). The interview questions were tailored to the roles of individuals and further refined throughout the research based on the findings of prior interviews. To complement this data, we used secondary data, generated by one of the abovementioned organizations, about uses of SM in self-management. The data consisted of eight interviews with people with long-term conditions who used SM for health purposes. Data from these interviews were analyzed together with the primary interview data.

### Table 1. Summary of primary data collection (X indicates inclusion).

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Description</th>
<th>Interview</th>
<th>Observation of online activities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Twitter</td>
</tr>
<tr>
<td>1</td>
<td>Organization 1</td>
<td>Professional sector</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>Organization 2</td>
<td>Intermediary</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>Organization 3</td>
<td>Intermediary</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>Organization 4</td>
<td>Intermediary, part of a larger project</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>5</td>
<td>Organization 5</td>
<td>Intermediary, part of a larger project</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>Alison Morgan</td>
<td>Project manager (FBG(^a) admin)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Sarah</td>
<td>Patient (forum and FBG user)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Edmund</td>
<td>Patient (forum, FBG, YouTube, and video blog user)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Carole</td>
<td>Patient (forum, FBG, and charity website user)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Donna</td>
<td>Patient and carer (Forum, FBG, and voluntary organization website user)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Karen</td>
<td>Carer (forum and FBG user)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Laura</td>
<td>Patient and carer (Forum and FBG user)</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)FBG: Facebook group.

### Table 2. Representation of the 4 constituent normalization process theory (NPT) constructs that attend to the 4 key aspects in eHealth implementation.

<table>
<thead>
<tr>
<th>NPT constructs</th>
<th>Coherence (sense-making work)</th>
<th>Cognitive participation (engagement or buy in work)</th>
<th>Collective action (enacting work)</th>
<th>Reflexive monitoring (appraisal work)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions</td>
<td>What gets done with social media (SM) in co-production? What gets done with other mechanisms?</td>
<td>How does SM facilitate participation within the intervention?</td>
<td>If an actor did not have SM, what would happen to his or her work (in terms of quality of service delivery)?</td>
<td>Does SM allow participants to reflect on the work they have done?</td>
</tr>
</tbody>
</table>
Finally, we observed the online activities of interview participants (organizational participants) and their uses of SM for health purposes. This enabled us to find evidence and complementary data to support the claims. Table 1 provides a complete list of the observation sources for each of the participants.

Data were collected over the period from March 2015 to December 2015. All conducted interviews were transcribed verbatim and transcripts checked for accuracy. We continued data collection until we judged that no new themes were identified and saturation was reached [24].

**Data Analysis**

Data were coded in NVivo software version 11 (QSR International) and thematically analyzed for each type of SM. We drew on the four principles of co-production (equality, diversity, accessibility, and reciprocity) as a deductive coding framework, extracting excerpts from our qualitative data that had bearing upon how SM reshapess co-production. In addition, we also inductively identified emerging themes surrounding the benefits and challenges of SM in enabling co-production in health and care, which served as an analytical lens to examine our data using a deductive approach to analysis [25]. Negative cases, that is, those that did not fit within the narrative, were explored in the most detail.

**Research Governance and Ethics**

This study was granted ethical approval by University of Edinburgh, School of Informatics. Consent forms were signed and agreed by all participating respondents. Identities were protected and assigned a confidential generic descriptor to ensure anonymity, and all names were changed.

**Table 3. Normalization process theory coding framework used for qualitative data analysis.**

<table>
<thead>
<tr>
<th>Coherence (sense-making work)</th>
<th>Cognitive participation (engagement or buy in work)</th>
<th>Collective action (enacting work)</th>
<th>Reflexive monitoring (appraisal work)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differentiation: What gets done in social media (SM)? What get done in other ways? What are the overlaps?</td>
<td>Enrollment: Can actors articulate the benefits of SM?</td>
<td>Skillset workability and Interactional workability: What do the users communicate through SM? To what extent does SM support co-productive work?</td>
<td>Reconfiguration: Do third party or charity organizations reflect on their activities on SM to develop new services through use of SM with co-production? Does reflection on SM contribute to redesign?</td>
</tr>
<tr>
<td>Communal Specification: How does SM contribute to the work? Do people agree with this as an account of the collaboration?</td>
<td>Activation: Can actors articulate how their work will change? Are they positive about this?</td>
<td>Contextual Integration and Relation al Integration: When users contribute in SM argument, does this have any influence on the decisions made? How does SM activity get captured and reused?</td>
<td>Communal Appraisal: How does SM influence coordination between organization and individuals in this context? Does SM let people build groups which are effective in service delivery?</td>
</tr>
<tr>
<td>Individual Specification: What does each actor use SM for? How is that different from what other actors do?</td>
<td>Initiation: Do actors understand their new activities involving SM and are they happy to conduct them?</td>
<td>Interactional Workability and Skillset Workability: How do responsibilities change?</td>
<td>Individual Appraisal: How do individual carers or service users appraise the effects of use of SM on them and their environment?</td>
</tr>
<tr>
<td>Meaning (internalization): What would be lost if SM were not used?</td>
<td>Legitimation: To what extent do actors and organizations believe that the action involving SM are important to the provision of the service?</td>
<td>Relational Integration and Contextual Integration: How does SM change the resource flow?</td>
<td>Systematization: How do organization (third sector or voluntarily) or individual users of SM in this context determine the effective (benefits or risks) or usefulness of SM in this context.</td>
</tr>
</tbody>
</table>

**Results**

Our findings show that overall existing SM helps support the four principles that underpin co-production—equality, diversity, accessibility, and reciprocity—and will influence the informal care sector to become more efficient. Below we explain how each principle of co-production can be enabled by existing SM. While appreciating the benefits, we also found tensions caused by use of SM as well as challenges that inhibit use of SM for co-production.

**Equality Through Sharing Experience of Users as Valued Assets**

To enable equality, individuals need to have the same status within a group and the group needs to recognize the value of the contribution of all individuals. Some types of SM (in particular, private FBGs) seem to allow recognition of skills and abilities of all members within a group.

Private FBGs were widely used by people who wanted to be connected to each other in a secure and closed manner. Participation in these groups needed to be approved by the administrator(s) based on whether individuals are patients or carers of a person with a particular condition. Therefore, those who were members of these groups held experiences, skills, or abilities in dealing with the condition. This knowledge was recognized by others as an asset that could be shared leading to a sense of being valued by others:

My knowledge is useful for others and their experience is valuable for me. We talk about our condition and liaising with each other and find ways to deal with issues...one particular case was when I
These experiences and skills either facilitated knowledge exchange or provided mental support, which in either case were seen as important to the group members. There are clear considerations of empowerment when people feel that their knowledge and skills are contributing to a change in the world. Although many positive consequences exist, we also need to be aware of the issues that may arise from this knowledge sharing and empowerment. These issues include the extent to which knowledge leads on to changes in the productivity of the health and care system (and possible lack of applicability of knowledge for some members of the group) and the means to prevent inaccurate or harmful information from propagating through the network. In similar terms, health and care professionals express concern over the unregulated transfer of experience through SM, which leads to a need for filtering and integration of information in such groups.

In many cases, the administrators of the groups also had the same condition as other members (or were the carers of people with the same condition). Having the condition meant that they were also equally concerned about the surrounding issues and had dealt with them for a considerable time. Thus, on the one hand, they brought comparable assets to the group, and on the other hand, they were equal in terms of status and position:

...with a closed group, you could have a moderator or an admin who works with that condition, so...they are going to actually facilitate the whole group, and without their, service provision, that group wouldn’t exist, and often the closed groups are not run by charities, they might just have been set up originally by someone who has had that particular experience, and they feel that there is a community for them of people in their situation out there, so they set it up themselves. [Patient]

As a result, although the member of these groups appreciated the equality of status, a new tension was created. Patients and carers acquired a considerable knowledge that could stand alongside health professionals; however, by no means were they equal in status or position to them. This in turn could lead to conflicts between the 2 groups.

This equality in terms of condition and experience removed the culture of “them and us” [26]. This, in turn, led to higher levels of support between all members (including administrators):

...they are volunteers who live with the condition, not employees of any organization. [Patient]

This was achieved by the closedness of the group (to ensure participants have similar levels of experience). However, this closedness could lead to tensions in terms of accessibility and diversity elements (discussed in the next section).

This equality in FBGs has empowered users to talk openly about their professional care practices and even discuss and find ways to approach professional carers (eg, general practitioners, National Health Service (NHS) consultants):

I definitely feel more in control too. For example, I was fobbed off a couple of years ago when asking a doctor for Vagifem and he said to use KY Jelly. The ladies here gave me the confidence to go back to my usual GP and ask assertively for the Vagifem I knew I needed. He agreed that Vagifem was a good idea and has prescribed it for me ever since. [Patient]

So, in general FBGs (and forums) generated a sense of community that facilitated equality among its users. However, there were times that things did not go as smoothly. Some members were aggressive about the stance they took on issues, which could lead to disagreement or, in more extreme cases, abandoning of the group:

Some people are militant when talking about their stances pro-anti surgery for Colitis and Crohns. They’ll really push their ideas on people and be very hard to talk to. You might have one person claiming to have the perfect solution to your problems: “Just cut out dairy!” Or someone else claiming that surgery or medication is a con by the health professionals. With Colitis and Crohns there are such extremes of symptoms and illness and a lot of people are frequently misdiagnosed due to this. [Patient]

SM Enables Diversity by Being Inclusive of Underrepresented Groups and by Connecting Diverse Groups of People

Diversity was enabled by SM in two ways. First, patients and carers are diverse in terms of characteristics (eg, literacy) and conditions. These differences can lead to less ability to access and use resources. Inclusiveness means overcoming these diversities and making sure that the people who are less likely to access or use resources are by some means gaining the benefit of these resources.

Patients mentioned that the closed nature of some SM, in particular, the private FBGs and forums, gave them the ability to talk about issues that cannot be discussed face-to-face because of embarrassment about conditions of particular illnesses. This meant that some of those who were formerly excluded because of their conditions could now benefit from these discussions:

People are more open about their experiences because it’s a closed group. They feel more open than if it was in the public domain...Online support takes away a lot of the social difficulties of sharing in a group for fear of embarrassment or sounding stupid. [Patient]

On a forum you talk about how you really feel, without any of the normal taboos. You can talk about anything. [Patient and Carer]

However, although this closeness of forums was an effective factor in facilitating some of these talks, it also created the challenge of getting into the groups. Thus, this closedness was a drawback as individuals could not join the groups without the permission of the administrators.

SM was not able to overcome many of the other barriers. For instance, interviewees highlighted that not everyone could have
access to various SM types such as FBGs and forums. This could be due to limited Internet access or low technological literacy.

Second, some types of SM, such as Twitter, acted as an effective place for connection of diverse people in health and care sector, including professionals and nonprofessionals (carer and patients). In comparison with many other SM, Twitter was used by a larger number of professional people:

*I think generally Twitter has certainly helped us to increase the amount of people that we have on the network. And also, to increase the amount of people that come along to the events. But again, we feel that that’s mostly in that professionals. So, we don’t really think that it’s been helpful in terms of targeting people with long-term conditions or carers at the moment.* [Organization participant]

As Twitter is a rapid and flat SM app, it provided a good space for users to find answers to their questions (without necessarily having to connect directly with people), getting current information and keeping up-to-date with health news:

*I think Twitter been used for exchange of informal information and really really useful information around about research. I found it extremely useful for the work on health literacies...So, you get to know people who are working and developing interesting stuff from health literacy...Twitter is good for following and that keep yourself up to date.* [Organization participant]

The flat nature of Twitter (no connections needed) also provided a good platform for raising funds or promoting campaigns by organizations and charities. In doing so, organizations used Twitter to promote their activities and keep all users updated. This is illustrated in the quotes below:

...so, we’d be very keen to promote our work [on Twitter], so we make sure that they’re linked to, we would be promoting. [Organization participant].

...it’s useful for campaigns as well, so there’s been a lot of really effective health campaigns on Twitter. [Organization participant]

However, issues, such as filtering imposed by the NHS in the use of SM on its premises, led to limitations in the use of such apps. One participant explained that their organization set up a blog; however, its use was constrained because of the firewall introduced by NHS that blocked access to blogs during daytime for professionals:

...there are a massive [number of] health care stuff using social media throughout our day but firewall is a big problem. [Organization participant]

So, although SM enabled diversity in terms of opening up a space for communication and knowledge sharing of some patients (and carers) with particular conditions, as well as offering a fast and flat platform for various actors (including health professionals, social workers, and carers) to share news, there were yet many barriers that limited the use of SM. As highlighted by the participants, individuals who had Internet accessibility issues could be excluded from gaining the benefit from SM. This could be either because of limited Internet access or the inability of some elderly people in using technology.

### SM Makes Groups’ Experience Accessible

To allow accessibility means everyone should have the same opportunity to participate in activities in a suitable manner. By offering various types of platforms (eg, blogs, FBGs, and Twitter), SM allowed different individuals to take part in knowledge sharing and communication in a way that suited them best:

*The one thing we found about Twitter, it seems to be very much used by the professionals. We find that most people with long-term conditions and carers will use Facebook. Whereas with Twitter, we will seem to target lots of professionals.* [Organization participant]

This allowed patients to gain access to some of the resources that were shared by professionals. Although it helps them reach a new layer of information about particular conditions, this did not mean having direct access to knowledge that leveraged their own condition. Therefore, accessibility was enhanced to some extent and for some of the users only.

Moreover, accessibility to group experience is enabled for those who have difficulty to gain access to others’ knowledge otherwise (such as through face-to-face meetings):

*I have quite a bad chest as you can hear, so I can be spending a lot of time on the forums or groups when I’m shut up in the house.* [Patient]

This accessibility to knowledge from various sources, in turn, empowered users, as illustrated by the quotes below:

*I would say that social media certainly empowers you. By people sharing their experiences, it makes you far more informed. You can find out what kind of treatments are out there and go to appointments armed with information. I also felt more empowered in how I dealt with health professionals if I felt I wasn’t being listened to. In fact, I later lodged a formal complaint to the health board.* [Patient]

*I’ve just had my results in from my test. GP, I saw him 2 times, never once told me that these results—and they were bad results. The GP missed it.* [Organization participant]

Although it increases patients’ knowledge, this was not necessarily welcomed by all professionals. Some professionals preferred to guide patients’ knowledge in certain directions. They believed that this knowledge is partial, and it will either lead to loss of trust or “interfere” with the course of their treatment (if patients take the advice from other sources rather than their direct healthcare professionals). They also believed that this knowledge does not take account of other issues such as limitations in NHS funding. Therefore, it can lead to new conflicts in terms of accessing scarce resources.

Another difficulty mentioned by patients was excessive online accessibility. This referred to the fact that sometimes too much online activity could lead to reduced physical activity. In more
Reciprocity SM Encourages Reciprocity in Sharing of Knowledge and Mental Support

Reciprocity refers to the mutual process of giving and receiving something back. Users of SM, in particular FBGs and forums, emphasized that they expected to gain something back from the group. Reciprocity may be direct (members behave in response to other members’ acts) or indirect (cooperation with strangers to gain reputation) [27]. Direct reciprocity could be generally seen in offering knowledge and experience about a topic:

Using social media is actually pretty empowering. When I was diagnosed, I had to become an expert on the condition and there’s no better source of knowledge for this condition than your own lived experience. I did a lot of personal research: first asking doctors and nurses about it, but the best information comes from the women who live with it. [Patient and Group administrator]

Indirect reciprocity, on the other hand, could be seen in offering mental support:

I wouldn’t want to join a group unless I thought that people would be able to empathize and understand what I’m going through. There’s no point in talking to people who don’t understand- they won’t respond appropriately. [Patient]

The sympathy that came from patients with similar health conditions (rather than paid organizational members) created added value for its recipients and led to the creation of a positive relationship:

The knowledge and information comes from the members of the group. It’s the people living with the condition who have the experience of self-managing, not paid employees of a charity who don’t necessarily live with a condition. [Patient]

Both forms of reciprocity played an important role in keeping the communities going. Therefore, administrators encouraged members to participate in talks, to make sure that everyone is receiving something back from the group.

We ask people to be active participants in the group: to commiserate with each other on a bad day, to be supportive of each other and share knowledge and experiences. [Patient and Group administrator]

Some administrators went further by deleting the members who were not active for a certain period of time:

People who don’t participate for more than a couple of months are deleted from the group. [Patient and Group administrator]

However, lack of involvement in discussions was sometimes due to lack of knowledge in the topic area or disagreement with the stance taken by other individuals. Therefore, administration of groups was a challenge:

Even if I don’t comment on posts, I read them so that I may be aware of any issues I may face...I don’t like the idea of taking HRT (Hormone Replacement Therapy) or any other things like creams and stuff. I prefer the natural route but I do understand now with information posted that each individual has their own opinions on the matter. These opinions and choices are personal to them and I take that on board now because this information is important knowledge. [Patient]

So although reciprocity was important in terms of the overall activities of individuals, the administrators needed to be considerate of members with lesser contributions. In some cases, some patients and carers started their participation as lurkers, just to get a feeling about the environment or to gain some specific knowledge. It would then take some time for them to reciprocate to the group. Therefore, user engagement could be seen as a gradual phased process. For those people with lower levels of engagement, who would be passive readers, it could begin by encouraging them to read more regularly, then starting to comment, and then contributing. The use of SM creates the opportunity to allow for growth of continuous knowledge and emotional conversation of strangers.

Discussion

Summary of Findings

This work indicates how different types of SM enable co-production by supporting its underlying principles: equality, diversity, reciprocity, and accessibility. The paper also offers insights into the challenges involved in use of these SM as an enabler of co-production. Individual users (patients and carers) and organizations providing health care services to elderly people adopted various kinds of SM to meet their diverse needs.
We observed that people’s contributions evolved as they became more experienced in the use of SM. Table 4 summarizes the benefits of each type of SM in terms of co-production principles.

In general, private FBGs were the most widely used SM by patients with similar conditions and their carers because of their greatest offerings around: (1) equality of members and valuing their experiences as assets; (2) diversity and inclusion of members whose voices are less heard otherwise; (3) accessibility for people from different geographical locations; and (4) reciprocity of knowledge sharing and mutual support. Forums were similar in terms of benefits and use; however, they were mainly sponsored (and administered) by organizations. This allowed for better control of data; however, their formation and access were more challenging. Microblogging (eg, Twitter) was also seen as one of the most highly used SM apps, which plays a very important role in health and care by both professionals (eg, doctors) and nonprofessionals (patients). Its “flat” nature allowed rapid exchange of information based on users’ interest in topics. Therefore, patients needing information or updates about particular diseases could easily gain access to information shared by health and care professionals. It was also highly used by those who wished to attract communities of interest or funds or those who wanted to provide or receive fast update about news and various topics. Therefore, it served for a very different purpose to those of FBGs and forums. Blogging, on the other hand, was used for slow but detailed sharing of stories by people and organizations about their health interests and experiences. We found four affordances of SM that supported care for elderly people: knowledge creation and sharing, information dissemination, emotional support, and new communication channels. SM afford behaviors that were difficult (or impossible) to achieve before these new tools were used by those involved in the care of elderly people. We further found mechanisms that affect how people engage in the knowledge and support conversation, which may have positive effects or may result in adverse consequences not intended by the participants or other groups involved in care of elderly people. These emergent tensions are the basis for the implications we draw.

In this way, SM offered new modes of communications not only between patients and their carers, but also between them and the professionals. On the one hand, professionals gained access to patient stories (blogs, FBGs, and Forums) and the details of conditions. This information can be used by doctors for better diagnosis and monitoring of particular patients. On the other hand, patients and carers gained access to new health and care findings.

Also, the joint effort in the creation of and monitoring of knowledge contents as well as the self-promoting nature of SM improved the productivity of health and care organizations by enabling them to publicize information using low-cost mediums.

**Interpreting Findings in the Context of the Wider Literature**

The large body of extant studies around the use of SM for health and care focus on who uses these tools [17,28-31] and uses of SM for communication [11,32-34]. The studies show that SM increases patients’ and carers’ access to health information [14,35-44]. Although our study confirms this, we specifically show that SM makes various types of health and care resources visible to meet the needs of elderly patients. These resources include availability of carers (including professional and nonprofessional resources), care programs (eg, outgoings, charity programs), knowledge about symptoms and cures of different conditions (including diets and drugs), new communication techniques with professionals, and more. We show that by facilitating new modes of dialog between different actors (ie, patient-patient, patient-carer, carer-carer, patient-professionals, and patient-healthcare organization), SM enables new, faster, and more effective modes of social interactions in which patients become empowered by having access to more resources.

SM offers a wide range of benefits for health communication, which can be grouped into increased interaction around general [17,45] and sensitive information [13], better accessibility of information [17,32,33,44,46-52], and emotional support [10,13,40,53-61]. We use Cohn’s co-production framework to expand the extant findings by showing how such characteristics act as the key principles of co-production. Our work shows that SM enables recognition of the experiences and skill of all participants as assets and enables them to engage with the community and become active.

**Table 4. Social media for co-production.**

<table>
<thead>
<tr>
<th>Co-production principle</th>
<th>Equality</th>
<th>Diversity</th>
<th>Accessibility</th>
<th>Reciprocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook groups</td>
<td>Patients with same condition and their carers; Experience and skills seen as asset</td>
<td>Less heard voices are included</td>
<td>Members from diverse geographical locations</td>
<td>Mutual support; knowledge sharing; administration of participation</td>
</tr>
<tr>
<td>Forums</td>
<td>Patients with same condition and their carers; Experience and skills seen as asset</td>
<td>Less heard voices are included</td>
<td>Members from diverse geographical locations</td>
<td>Mutual support; knowledge sharing; administration of participation</td>
</tr>
<tr>
<td>Microblogging</td>
<td>Professionals and nonprofessionals; No direct connections needed</td>
<td>Retweets; provides access to another social media</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blogging</td>
<td>Accessible by all</td>
<td></td>
<td></td>
<td>Feedback on blogs</td>
</tr>
</tbody>
</table>
Our study also expands the existing literature, by showing that the fulfilling of different needs by various SM is influenced by different factors, including the speed of knowledge creation and dissemination, the speed of feedback and discussions, the detailed nature of knowledge exchanges, the type of discussion (support vs news vs health knowledge sharing), and the openness and closedness of activities. These characteristics help better coordination and communication of knowledge resources between carers and patients.

There are also limitations in the use of SM. Information quality concerns and the lack of reliability of the health information [10,38,40,41,45,53,62-69] are among the widely discussed limitations. Although our findings confirm these, we also show that the explication of common resources may lead to new forms of competition and conflicts. In particular, the new knowledge that is obtained by users is not always welcomed by professionals. This could be due to numerous reasons, including lack of validity of all information obtained, as well as higher demand for treatments as they become known to patients and carers. Also, because of concerns about information quality and validity, some health care organizations need to put into place new forms of information monitoring, which may be costly.

Strengths and Limitations

This paper has a number of strengths and limitations. We drew on NPT [23,70], which served as a sociotechnical analytical lens, to help us analyze the benefits as well as challenges of various types of SM. We have drawn data from multiple different sources, including patients, carers, and charity organizations to enhance confidence in our findings and included diverse perspectives. However, because of the sensitivity of patient data, we only had limited access to private FBGs and forums. We overcame this problem by contacting many groups and gaining access to one particular group. To also understand other groups that were important for this research, instead of observations, we interviewed its users. We also did not seek the perspective of NHS professionals including doctors. This can be addressed in future research with a focus on professionals. Finally, in this paper, we have focused on communication and cooperative utilization of health and care resources. Therefore, further research is needed to focus on resourcing, conflict and competition for resource, and the overall governance of health and care provision.

Conclusions

SM has gained momentum within the health and care community by offering significant benefits for patients, carers and even professionals; increasing interaction; providing more readily available and customized information; offering mental support; promoting health and care–related activities; offering a platform for communication for underrepresented individuals; allowing reciprocal sharing; and enhancing the communication between patients, carers, and professionals. All these benefits have the potential to be realized through SM. These benefits facilitate co-production by enhancing equality, diversity, accessibility, and reciprocity, and lead to recognition of resources (skills and time), joint creation and monitoring of knowledge, and direct and indirect mutual support. This in turn can lead to resource savings needed to manage the growth in demand from the expanding elderly population. SM allows users to learn from each other (in a less costly manner) and can facilitate communication more effectively (in particular, professionals and nonprofessionals).

However, despite these benefits in facilitating co-production, existing SM does not fully enable co-production. There are as yet outstanding issues in arranging the common pool of health and care resources to better enable co-production. Different SM enable co-production (co-delivery) of services for elderly people to varying extents. In particular, SM is used distinctly differently by professionals and nonprofessionals. This can be seen as an opportunity to leverage their benefits in a more productive manner.

Acknowledgments

We gratefully acknowledge the individuals and organizations who agreed to participate in this study. In the course of gathering this material, interviewing the participants, and collecting the data, we had invaluable inputs from those who volunteered to share their life stories.

Authors' Contributions

HD and SA conceived this work. HD collected data for this study. HD led on data analysis and drafting of the manuscript. All authors (HD, SA, HM, and RW) have commented on various versions of this manuscript and inputted into the analysis.

Conflicts of Interest

None declared.

References


Abbreviations
- FBG: Facebook group
- ICT: information and communications technology
- NPT: normalization process theory
- SM: social media
- NHS: National Health Service
VA FitHeart, a Mobile App for Cardiac Rehabilitation: Usability Study

Alexis L Beatty¹,², MAS, MD; Sara L Magnusson¹,², BA; John C Fortney¹,³, PhD; George G Sayre¹, MA, PsyD; Mary A Whooley⁴,⁵,⁶, MD

¹Veterans Affairs Puget Sound Health Care System, Seattle, WA, United States
²Department of Medicine, University of Washington, Seattle, WA, United States
³Department of Psychiatry, University of Washington, Seattle, WA, United States
⁴San Francisco Veterans Affairs Medical Center, San Francisco, CA, United States
⁵Department of Medicine, University of California, San Francisco, San Francisco, CA, United States
⁶Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Alexis L Beatty, MAS, MD
Veterans Affairs Puget Sound Health Care System
1600 S Columbian Way
Seattle, WA,
United States
Phone: 1 206 764 2008
Email: beattya@uw.edu

Abstract

Background: Cardiac rehabilitation (CR) improves outcomes for patients with ischemic heart disease or heart failure but is underused. New strategies to improve access to and engagement in CR are needed. There is considerable interest in technology-facilitated home CR. However, little is known about patient acceptance and use of mobile technology for CR.

Objective: The aim of this study was to develop a mobile app for technology-facilitated home CR and seek to determine its usability.

Methods: We recruited patients eligible for CR who had access to a mobile phone, tablet, or computer with Internet access. The mobile app includes physical activity goal setting, logs for tracking physical activity and health metrics (eg, weight, blood pressure, and mood), health education, reminders, and feedback. Study staff demonstrated the mobile app to participants in person and then observed participants completing prespecified tasks with the mobile app. Participants completed the System Usability Scale (SUS, 0-100), rated likelihood to use the mobile app (0-100), questionnaires on mobile app use, and participated in a semistructured interview. The Unified Theory of Acceptance and Use of Technology and the Theory of Planned Behavior informed the analysis. On the basis of participant feedback, we made iterative revisions to the mobile app between users.

Results: We conducted usability testing in 13 participants. The first version of the mobile app was used by the first 5 participants, and revised versions were used by the final 8 participants. From the first version to revised versions, task completion success rate improved from 44% (11/25 tasks) to 78% (31/40 tasks; P=.05), SUS improved from 54 to 76 (P=.04; scale 0-100, with 100 being the best usability), and self-reported likelihood of use remained high at 76 and 87 (P=.30; scale 0-100, with 100 being the highest likelihood). In interviews, patients expressed interest in tracking health measures ("I think it’ll be good to track my exercise and to see what I’m doing"), a desire for introductory training ("Initially, training with a technical person, instead of me relying on myself"), and an expectation for sharing data with providers ("It would also be helpful to share with my doctor, it just being a matter of clicking a button and sharing it with my doctor").

Conclusions: With participant feedback and iterative revisions, we significantly improved the usability of a mobile app for CR. Patient expectations for using a mobile app for CR include tracking health metrics, introductory training, and sharing data with providers. Iterative mixed-method evaluation may be useful for improving the usability of health technology.

(JMIR Hum Factors 2018;5(1):e3) doi:10.2196/humanfactors.8017

http://humanfactors.jmir.org/2018/1/e3/
KEYWORDS
cardiac rehabilitation; mobile applications; exercise therapy; exercise; rehabilitation research; telemedicine; habits; qualitative research

Introduction

Cardiac rehabilitation (CR) is an evidence-based program of exercise training, risk factor management, education, and counseling that improves outcomes for patients with heart disease [1-4]. However, CR is dramatically underused, with less than 20% of eligible patients participating [5-7]. Many barriers limit participation, including expectations for attending facility-based supervised exercise sessions three times per week for 12 weeks, transportation difficulties, competing demands related to work or family, lack of social support, and cost [8-10]. Home-based CR programs are similar in efficacy and safety to facility-based programs but have not been widely adopted in the United States [11,12]. New strategies are needed to promote participation in home-based CR [13,14].

Technology has the potential to facilitate health interventions and motivate patients to improve health behaviors, including for the secondary prevention of cardiovascular disease [15-19]. It is known that interventions with a theoretical basis are more effective [20], but most technology solutions have not been created around evidence-based practices or health behavior theory [18,21-23]. The Theory of Planned Behavior (TPB) [24] has been successfully applied to CR in both facility- and home-based settings [25,26]. The TPB states that the most important determinant of behavior is the intention to perform the behavior. Behavioral intention is influenced by constructs of attitudes, subjective norms, and perceived behavioral control. An extension of the TPB has been developed to explain behavior specific to technology use, called the Unified Theory of Acceptance and Use of Technology (UTAUT) [27] and its extension for consumer use of technology (UTAUT2) [28]. This theory contends that constructs of performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit influence behavioral intention, which is the strongest predictor of technology use.

Using the TPB and UTAUT2, we developed a theory-based mobile app for technology-facilitated home CR. We tested the mobile app in patients eligible for CR, obtained feedback, and iteratively made revisions to the mobile app to improve its usability. Additionally, we interviewed participants about physical activity, CR, and mobile app use to better understand how to implement technology-facilitated home CR. The aims of this study were to determine the usability of the VA FitHeart mobile app and to analyze factors contributing to its use.

Methods

Overview

We conducted an observational study of Veteran use of a mobile Web app, VA FitHeart. The mobile app was designed to be used as a tool for home CR and includes physical activity goal setting, logs for physical activity and health measures (eg, blood pressure, pulse, weight, glucose, cholesterol, and mood), health education, reminders, and feedback (Figure 1). The mobile app was developed by the Department of Veterans Affairs (VA), and testing was conducted on versions of the mobile app hosted on preproduction testing servers. VA FitHeart was designed with input from subject matter experts (including the authors), patients eligible for CR, user experience designers, and mobile app developers. VA FitHeart underwent iterative user interface revisions based on participant feedback on usability. Because the app was hosted in a testing environment, there were occasional server downtimes when the app was not accessible for testing.

Participants

Veterans attending the outpatient cardiology clinic at the VA Puget Sound Health Care System in Seattle, WA were screened for enrollment in the study. Eligibility criteria included the ability to speak English, age ≥21, and eligibility for CR, defined as myocardial infarction, percutaneous coronary intervention, or cardiac surgery in the past year or having chronic stable angina or heart failure. Participants were excluded if they were not eligible for CR. Participants meeting inclusion criteria were asked to participate in additional screening to participate in a study about a mobile app for CR. Participants were excluded if they did not have access to a mobile phone, tablet, or computer with Internet access. This study was reviewed and approved by the institutional review board at the VA Puget Sound Health Care System. All participants provided written, informed consent.

Usability Testing

Study staff demonstrated the mobile app to participants in person and asked participants to complete prespecified tasks with the mobile app while study staff observed the participants. Tasks demonstrated by study staff included setting a physical activity goal, making a physical activity entry, viewing a fitness graph, making a weight entry, and viewing an educational module. After the conclusion of the demonstration, participants were asked to complete the demonstrated tasks independently. Study staff recorded task completion success if the participant was able to successfully complete the task.

Questionnaires

Following testing, participants completed questionnaires using REDCap electronic data capture tools hosted at the VA [29], including rating their likelihood to use the mobile app from 0 (low) to 100 (high). Participants completed the System Usability Scale (SUS) [30], with scoring from 0 to 100, with ratings of greater than 70 generally considered to demonstrate acceptable usability [31]. In addition, participants rated factors influencing mobile app use related to constructs from UTAUT2, including performance expectancy, social influence, facilitating conditions,
habit, hedonic motivation, price value, and behavioral intention (scale 0-100; Multimedia Appendix 1) [28]. Effort expectancy was operationalized as response to the SUS.

**Interviews**

We conducted two separate semistructured interviews with Veterans enrolled in the study. The first interview was conducted before usability testing and was centered on physical activity and the use of technology. The second interview was conducted before usability testing, asking specific questions about the functionality of the mobile app. All interviews took place in person at the VA Puget Sound Health Care System in Seattle, WA in a private office. Both interviews had semistructured interview guides that included open-ended questions and prompts for elicitation of additional detail (Multimedia Appendix 2). Interviews were conducted by two trained study staff members, audiorecorded, and transcribed word for word.

**Figure 1.** Screenshots of VA FitHeart, a mobile app for cardiac rehabilitation.
Analysis

Descriptive statistics of range and mean were used for quantitative questionnaire responses. To compare responses before and after, we performed a two-tailed \( t \) test. Qualitative interviews were transcribed and coded using Atlas.ti (Atlas.ti GmbH) version 7.2. Two researchers (ALB and SLM) coded interviews; we performed inductive and deductive content analysis [32]. We used a priori categories from the constructs of the UTAUT for consumer applications (performance expectancy, effort expectancy, social influence, facilitating conditions, habit, hedonic motivation, and price value) and the TPB (attitudes, subjective norms, and perceived behavioral control). In addition, we generated additional codes that emerged naturally based on participant responses. Both researchers wrote analytic memos to document observations and participated in intermittent meetings to discuss emergent themes, add or collapse codes, and reach consensus on coding disagreements. The research team conducted a thematic analysis to assess patterns of experiences and opinions across themes and reached agreement of interpretation. Analysis was conducted concurrently with participant enrollment. We continued enrollment of new participants until we had achieved acceptable usability and stakeholders believed that sufficient data had been collected to make the decision to not make additional revisions. The results of the study are reported in accordance with the Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth checklist [33].

Results

Participant Characteristics

From January 27, 2016 to October 24, 2016, we enrolled 15 participants in usability testing (Multimedia Appendix 3). Participants ranged in age from 43 to 75 years (mean 63 years). There were 14 males and 1 female, and 13 participants identified race as white (87%). Primary diagnoses included coronary artery bypass surgery (2/15, 13%), percutaneous coronary intervention (3/15, 20%), chronic stable angina (5/15, 33%), and stable heart failure (6/15, 40%).

Usability Testing

The first version of the mobile app was used by the first 5 participants, and revised versions were used by 8 participants. Two participants were unable to complete testing because of technical difficulties with accessing the servers in the preproduction testing environment during server downtimes. From the first version to revised versions, task completion success rate improved from 44% (11/25 tasks) to 78% (31/40 tasks; \( P=.05 \)), SUS improved from 54 to 76 (\( P=.04 \); scale 0 to 100, with 100 being the best usability), and rated likelihood of using the mobile app remained high at 76 and 87 (\( P=.30 \); scale 0 to 100, with 100 being the highest likelihood; Figure 2). We found that revised versions of the mobile app significantly improved constructs from UTAUT2, including effort expectancy, habit, and hedonic motivation (Figure 3).

Figure 2. Task completion success and patient-reported usability and likelihood of using the mobile app on initial and revised versions of the mobile app. Task completion success was the percentage of tasks successfully completed. Usability was score on the System Usability Scale (scale 0-100, with 100 being the best usability). Likely to use app was self-rated likelihood of use (scale 0-100, with 100 being the highest likelihood). \( P \) values represent comparisons between the initial version and revised versions.
Figure 3. Patient-reported factors influencing mobile app use on initial and revised versions of the mobile app. Items were rated on a scale of 0 to 100, with 100 being the best rating. *P<.05 for comparison between versions.

Mobile Technology Use
Emergent themes about mobile technology use were categorized by UTAUT2 construct (Multimedia Appendix 4).

Performance Expectancy
Many participants expect that VA FitHeart would be beneficial.

*I think that the idea of an app that records all of the information that this app is doing will be very valuable. Actually somewhat of a motivation for me to do this thing.* [P28]

Participants desired that a mobile app for CR be able to track goals, physical activity, and other health measures such as blood pressure, heart rate, weight, blood glucose, and diet.

Although there were suggestions for additional features to the mobile app, such as the ability to integrate with sensors and automatically transfer data, it was commented that this was not essential.

*Memorizing, writing it down and then getting it into your computer, if that was all done while you’re doing activities and stuff that would be a big help. But if they can’t, this is still a good app. Still helpful.* [P28]

Effort Expectancy
Several aspects of ease of use of the mobile app emerged. Participants appreciated simplicity.

*It was pretty easy...I like that it’s simple.* [P45]

The flow is very simple. [P07]

Vision and size of text were cited as a barrier by many participants.

*The only downside I see for me is with my vision; the fonts are a little small. I would definitely need to use my reading glasses to read it.* [P44]

Prominent display of key features was cited as a facilitator of ease of use.

*The settings to change your goals are very easy to reach and very prominent.* [P23]

Although some users commented on functions that were not as intuitive and harder to find, it was recognized that with more experience and familiarity, this problem could be overcome.

*I’m not used to this. Once I get used to it, I’ll know where everything is.* [P40]

One general barrier to ease of use mentioned by participants was the use of passwords and codes. This did not emerge as a barrier specific to our app, but participants were not required to enter a password during the testing session.

Social Influence
Participants often mentioned a desire to share their data with their providers.

*I like the fact that I can put all of that and track it, and that my doctors can as well. I can show my doctor what I’ve been working on.* [P45]

There was also interest in communicating with providers through the app. Family and peer support were reported to influence mobile technology use. The mobile app does feature a link to an online social networking site for patients with heart disease, but social networking was infrequently mentioned.

Facilitating Conditions
A desire for hands-on initial training on how to use the mobile app emerged as an important theme.
Initially, training with a technical person, instead of me relying on myself. [P8]

Expectations for additional help varied, including online, telephone, and family or peer support.

If I had problems I’d try to find out how to fix it on this or call you. [P40]

But I’ve got 3 boys that are all pretty much wizards at it, but I’m not. I’m sure I can learn it or if they punch in the application so that it could come right up. I’d be fine. [P19]

Habit

Habit was frequently mentioned by participants, both with regard to their use of technology and related to participating in physical activity. Habit was also linked by many Veterans to their previous military service. Our interview guides did not specifically probe participants about habit, making the prominent emergence of habit notable. In the discussion of habit, some participants described how memory and learning contribute to the development of habitual use of technology.

Memory appears to play a dual role in use of the technology—in remembering to use the technology and how to use the technology.

Something to remind me. But, I’m going to have to set a schedule of when I actually do this. [P13]

It’s a problem with my memory. The program to me seems fine if I can remember how to go through it. [P15]

Learning was discussed often as a period of trial and error where users would become more facile with using the app with greater experience.

Once I learned this app and spent just a little bit of time with it, I’ll be good with it. I don’t see any problem with it. [P23]

Ultimately, these efforts are expected to result in habitual use of VA FitHeart.

If I were to [use the app] religiously, every day do it, then it’d be force of habit. [P08]

Hedonic Motivation

Most comments about pleasure derived from using technology were general in nature. Comments about VA FitHeart itself were less strongly pleasurable in nature, but generally positive.

But I like the looks of the app and I like what it’s set up to do. [P28]

Price Value

Though participants mentioned price and cost related to other technologies and mobile apps, price value was infrequently mentioned linked to our mobile app, which will be free for general use.

I think in the end, you could save people, or patients, money. [P35]

Physical Activity

In our interviews, we identified many of the common barriers and facilitators to physical activity and participation in CR that have been described in previous studies (Multimedia Appendix 5) [34].

Attitudes expressed included general attitudes toward physical activity, as well as comments related to health benefits and the influence of other medical conditions. Many participants commented on subjective norms including the influence of pets, family, and health care providers. Participants frequently mentioned themes relating to perceived behavior control such as goals, habit, motivation, work (as either a facilitator or barrier), and travel or transportation.

We identified one notable emergent theme that does not clearly fall within a single TPB construct and that has not been well described before: the role of military service in physical activity.

Military Service

Though we specified a priori categories, the topic of military service was mentioned so frequently by our population that we created an emergent category for military service, which may be uniquely important to our patient population. In our population of US military Veterans, almost all Veterans reported their time of military service as a physically active time in life. Their time in military service was often central to their experience related to physical activity.

When I joined the service I was very fit. I usually did physical activity in the morning and sometimes in the afternoon also, an average of 2.5 hours a day, 4 to 5 days a week. [P7]

Additionally, many Veterans described their time after discharge as a particularly inactive time.

I hadn’t worked out since the military. It had been like 18 years since I’d set foot in a gym. [P45]

Discussion

Principal Findings

We found that iteratively revising a mobile app for CR based on user feedback resulted in significant improvements in the usability of the mobile app. Using a theory-based approach, we revealed interest in using a mobile app to track physical activity and health measures and to share data with providers. Patients expected to have training on how to use the mobile app. On the basis of participant comments, establishing habit, both with regard to physical activity and mobile app use, is anticipated to be a key contributor to adoption of this technology.

This is the first theory-based investigation of the usability of a mobile app for CR. It is known that interventions based on theory are more likely to be effective [20]. Other technology-facilitated interventions for CR have been studied, with promising results [18,35,36]. However, these studies did not describe theoretical considerations related to health behaviors or technology use, so we know little about how the interventions influenced patient behavior to achieve their results. Other investigators have also reported the development of
theory-based mobile CR platforms, but results of their use and efficacy have not been reported [37]. Having a framework for understanding how an intervention produces its effects will be important for studying its impact and adapting interventions beyond research studies. We found that constructs from UTAUT2 [28], especially performance expectancy, effort expectancy, social influence, facilitating conditions, and habit appear to play an important role in use of mobile technology for CR.

Patients in our study desired the ability to track physical activity and health measures with an easy-to-use mobile app, confirming findings from previous studies [18,38,39]. Though some participants expressed a desire for additional features to the mobile app, such as integration with device or peripheral sensors for motion or location, it was commented that these features were not essential. In general, VA FitHeart received praise for its simplicity.

It has previously been reported that people have little desire to share their personal fitness data with their providers [40]. We found that many patients expected to share their data with their health care providers and viewed this as a key advantage to using VA FitHeart. It may be that apps designed to be used for health conditions are viewed differently than consumer personal fitness trackers. Other studies of patient-provider digital communication interventions have demonstrated high levels of satisfaction [41]. Theory related to physical activity behavior and technology use behavior would suggest that sharing data with providers has the potential to influence patient use of a mobile app to promote physical activity through subjective norms and social influence [24,28], and our finding that patients expect to share their data with providers is consistent with this.

Many participants expressed an expectation for in-person training on use of the mobile app, in addition to on-demand help online, via telephone, or from family and friends. Previous studies of older adults have also revealed a preference for in-person training and the influence of family and friends [42,43]. It has also been suggested that technology training for older adults may need to be geared toward their needs and learning styles [44,45]. As older adults are less likely to use mobile technology than younger adults, interventions and training geared toward older adults may be necessary [46]. Together, this suggests that interventions for technology-facilitated CR should include opportunities for in-person training of participants on use of the technology, in addition to on-demand help.

Habit was frequently and prominently mentioned by Veterans as a factor that will be important, both for using the mobile app and participating in physical activity. UTAUT2 describes experience and habit as related concepts, with experience being necessary but not sufficient for establishment of habit [28]. In our study, patients frequently discussed memory and learning as prerequisites to habitual use, rather than mentioning experience. For our older population, experience may need to be considered more broadly with regard to repetition and retention of learned behaviors to establish habitual use. In addition, mention of habit was linked by some Veterans to their military service, and it is possible that experience in military service influences how habit is developed in our population. Interestingly, with iterative revisions to improve the usability of the mobile app, we noticed improvements in participant ratings of expected habit and hedonic motivation with use. Though effort expectancy is not theorized to influence habit or hedonic motivation [28], it may be that the usability of a mobile app influences expected adoption of regular use and pleasure derived from the mobile app. Other studies have found that for new users of online fitness communities, self-regulatory motives influence habitual use but that for experienced users, social motives and enjoyment play a larger role [47]. It has also been observed that for social apps, perceived usefulness and hedonic motivation influence habit, which may mediate the effects of perceived usefulness and hedonic motivation on technology use behavior [48]. Together, this suggests that mobile apps that are easier to use may be both more enjoyable to use and more likely to be perceived as habit-forming.

Limitations
Several limitations to our findings should be considered. We had a small sample size of Veterans and only one female, so our population may not fully represent the population or non-Veteran populations. As not all eligible patients agreed to participate, our findings may not be representative of the entire eligible population. Due to our small sample size, we may not have truly achieved thematic saturation of all factors associated with the use of mobile technology for CR. However, our sample did provide valuable feedback that resulted in improved usability. Additionally, we studied VA FitHeart in a usability testing environment and not in a real-world environment. Further testing is needed in a real-world environment to determine whether other factors are important to use.

Conclusions
With participant feedback and iterative revisions, we significantly improved the usability of a mobile app for CR. Patient expectations for using a mobile app for CR include tracking health metrics, introductory training, and sharing data with providers. Iterative theory-based mixed-method evaluation may be useful for improving the usability of health technology.

Acknowledgments
The authors appreciate the support of the Veterans Health Administration Office of Connected Care, Veterans Health Administration National Cardiology Office, Veterans Health Administration Office of Rural Health, and Veterans Affairs Office of Information and Technology in the development of VA FitHeart.
Conflicts of Interest

This study was funded by the US Department of Veterans Affairs Virtual Specialty Care Quality Enhancement Research Initiative (QUE 15-282), the Seattle Institute for Biomedical and Clinical Research, the John L. Locke Jr. Charitable Trust, and the Alpha Phi Foundation. ALB is supported by Career Development Award #16-5150 from the VA Health Services Research and Development Program. The views expressed in this study are those of the authors and do not necessarily reflect the position or policy of the VA or the US Government.

Multimedia Appendix 1
Mobile Application Use questionnaire.

[PDF File (Adobe PDF File), 33KB - humanfactors_v51e3_app1.pdf ]

Multimedia Appendix 2
Interview guides.

[PDF File (Adobe PDF File), 41KB - humanfactors_v51e3_app2.pdf ]

Multimedia Appendix 3
Screening and enrollment of participants.

[PNG File, 16KB - humanfactors_v51e3_app3.png ]

Multimedia Appendix 4
Quotations related to concepts from the extended unified theory of acceptance and use of technology.

[PDF File (Adobe PDF File), 74KB - humanfactors_v51e3_app4.pdf ]

Multimedia Appendix 5
Quotations related to concepts from the Theory of Planned Behavior.

[PDF File (Adobe PDF File), 72KB - humanfactors_v51e3_app5.pdf ]

References

1. Balady GI, Williams MA, Ades PA, Bittner V, Comoss P, Foody JM, American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee, the Council on Clinical Cardiology, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Epidemiology and Prevention, American Heart Association Council on Nutrition, Physical Activity, and Metabolism, American Association of Cardiovascular and Pulmonary Rehabilitation. Core components of cardiac rehabilitation/secondary prevention programs: 2007 update: a scientific statement from the American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee, the Council on Clinical Cardiology; the Councils on Cardiovascular Nursing, Epidemiology and Prevention, and Nutrition, Physical Activity, and Metabolism; and the American Association of Cardiovascular and Pulmonary Rehabilitation. Circulation 2007 May 22;115(20):2675-2682 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.106.180945] [Medline: 17513578]


31. Sauro J. Measuring U. Measuring usability with the system usability scale (SUS) URL: https://measuringu.com/sus/ [WebCite Cache ID 6qMTV09ya]


---

**Abbreviations**

- **CR:** cardiac rehabilitation
- **SUS:** System Usability Scale
- **TPB:** Theory of Planned Behavior
- **UTAUT:** Unified Theory of Acceptance and Use of Technology
- **UTAUT2:** Unified Theory of Acceptance and Use of Technology extension for consumer use of technology
- **VA:** Veterans Affairs
How Health Care Professionals Evaluate a Digital Intervention to Improve Medication Adherence: Qualitative Exploratory Study

Karen Thomson, Corline Brouwers, Olga C Damman, Martine C de Bruijne, Danielle RM Timmermans, Marijke Melles

Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands
Department of Public and Occupational Health and Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, Netherlands

Corresponding Author:
Corline Brouwers, PhD
Department of Public and Occupational Health and Amsterdam Public Health Research Institute
VU University Medical Center
Van der Boechorststraat 7
Amsterdam, 1081 BT
Netherlands
Phone: 31 204448263
Email: c.brouwers@vumc.nl

Abstract

Background: Medication nonadherence poses a serious and a hard-to-tackle problem for many chronic diseases. Electronic health (eHealth) apps that foster patient engagement and shared decision making (SDM) may be a novel approach to improve medication adherence.

Objective: The aim of this study was to investigate the perspective of health care professionals regarding a newly developed digital app aimed to improve medication adherence. Familial hypercholesterolemia (FH) was chosen as a case example.

Methods: A Web-based prototype of the eHealth app—MIK—was codesigned with patients and health care professionals. After user tests with patients, we performed semistructured interviews and user tests with 12 physicians from 6 different hospitals to examine how the functionalities offered by MIK could assist physicians in their consultation and how they could be integrated into daily clinical practice. Qualitative thematic analysis was used to identify themes that covered the physicians' evaluations.

Results: On the basis of the interview data, 3 themes were identified, which were (1) perceived impact on patient-physician collaboration; (2) perceived impact on the patient's understanding and self-management regarding medication adherence; and (3) perceived impact on clinical decisions and workflow.

Conclusions: The eHealth app MIK seems to have the potential to improve the consultation between the patient and the physician in terms of collaboration and patient engagement. The impact of eHealth apps based on the concept of SDM for improving medication-taking behavior and clinical outcomes is yet to be evaluated. Insights will be useful for further development of eHealth apps aimed at improving self-management by means of patient engagement and SDM.

(JMIR Hum Factors 2018;5(1):e7) doi:10.2196/humanfactors.8948

KEYWORDS
medication adherence; eHealth; shared decision making; self-management; patient engagement

Introduction

Medication Nonadherence

Medication nonadherence is a major problem faced by people with chronic conditions [1]. Nonadherence can occur both unintentionally (due to a lack of capacity or resources; eg, poor memory) and intentionally (active decision of the patient; eg, due to medication intolerance) [1,2]. The outcomes of nonadherence are well-known—loss of opportunities for patients to improve their health and the loss of medication by health care systems, with the subsequent effect of increased morbidity [3]. Identifying the principal causes of nonadherence has proven to be complex [4,5]. Medication nonadherence neither seems to be directly related to the type or severity of a disease [6] nor to individual traits or sociodemographic characteristics [7,8]. Patients' knowledge, beliefs, and concerns regarding treatment, as well as their actual experiences with side effects do seem to
be essential factors influencing medication adherence, especially *intentional* nonadherence [8-11].

**Interventions to Improve Medication Adherence**

In recent years, many interventions have been developed to improve medication adherence, but these are often insufficiently successful or effective [12,13]. This is particularly seen in short-term interventions, such as counseling, written information, and personal phone calls. Long-term interventions with multiple components (eg, more convenient care, information, counseling, reminders, self-monitoring, reinforcement, family therapy, psychological therapy, mailed communications) are, in general, more likely to show benefits. However, these interventions often show a disproportional distribution between the benefits on one hand and the high expenditure of time by health care professionals and (consequent) financial resources on the other hand. Hence, there is a growing interest in digital interventions that could be time-saving [1,14]. Most digital interventions (apps), currently available for medication adherence, have functionalities such as medication reminders, medication diaries, and access to medication instructions. These apps are mainly focused on nonintentional adherence [15], and they are usually targeted only to the individual patient rather than the interaction between patients and their health care professionals. As a consequence, the health care professional’s opportunity to support patients in improving adherence may not be optimally utilized.

To target intentional nonadherence, we developed a digital intervention in collaboration with patients and health care professionals, which focused on the patient’s preferences and beliefs about treatment options and on their actual experienced side effects and quality of life. The intervention was designed to foster patient engagement (thus, medication adherence) using 2 routes: (1) prompting patients and professionals to be aware of and discuss the patient’s preferences and beliefs about his or her current health and treatment regimen in the consultation, which is based on the Necessity–Concerns Framework (NCF) and the models of shared decision making (SDM) [8,16] and (2) increasing the patients’ engagement with management of their disease outside the consultation, through enhancing their knowledge and insight into their health status over time in relation to the medication/treatment regimen (a self-management approach) (Figure 1) [17]. By explicitly discussing patients’ beliefs, preferences, and concerns in the consultation, it seems more likely that physicians and patients choose a treatment regimen that is adhered to by the patient. Moreover, such engagement is also likely to ensure that patients take more responsibility for their health and promptly contact their physician when they encounter problems with their medication.

**Case Study: Familial Hypercholesterolemia**

The genetic condition familial hypercholesterolemia (FH) was chosen as a case example for developing the digital intervention. FH patients have increased levels of low-density lipoproteins, which makes them prone for developing cardiovascular diseases (CVDs). Current estimations suggest that 1 out of every 240 people have FH [18]. Clinical guidelines state that statin medication should be the cornerstone in the treatment of FH patients [19]. In addition to statins, a considerable number of patients also need other types of lipid-lowering medications to reach optimal treatment effects (ie, reduction in the level of low-density lipoprotein). Within this regimen of lipid-lowering medications, there are decisions to be made about the type and dosage of the medications. Apart from medication, FH patients are always advised to adopt a healthy lifestyle [19].

The overall medication adherence in FH patients ranges between 58% and 89% [20,21], indicating that a substantial number of patients are nonadherent. So far, the current literature has failed to adequately explain nonadherence among FH patients [7,21].

**Figure 1.** Schematic overview of functions of MIK. LLT: lipid lowering therapy; QoL: quality of life; LDL: low density lipoprotein.
As for other conditions, it is likely that adherence problems among FH patients are caused by an interplay of factors relating to patients’ beliefs, values, and experiences with side effects [1,22], as well as by factors relating to patient-professional communication [3,9,23-25]. It could be that FH patients experience a low sense of urgency because they typically do not (yet) experience actual health complaints because of FH. Additionally, the medication regimen for FH patients is lifelong and finding the right medication for FH patients is often a trial-and-error process, making medication adherence a challenge.

The developed digital app aimed to improve medication adherence of FH patients was named “MIK” (Dutch for “to aim”). After a participative human-centered design (HCD) process, involving both patients and health professionals, the final concept of MIK was first evaluated by FH patients in a pilot test. FH patients highly valued the fact that they were being triggered to think about their preferences regarding treatment and topics that they would like to discuss with their health care professional. More importantly, patients mentioned that MIK would improve their sense of control by providing an overview of important data and provide an opportunity to change their conversation with the health care professionals. The aim of this study was to investigate the perspective of health care professionals regarding MIK. The user tests and semi-structured interviews addressed: (1) whether the designed functionalities aimed at improving medication adherence fit the needs of health professionals; (2) how health professionals would use and interpret the information provided by MIK; and (3) what barriers and facilitators for the use of MIK in daily practice were identified by health professionals.

Methods

A qualitative explorative evaluation study was conducted among health care professionals to investigate their perspective regarding the designed set of eHealth functionalities in MIK, which aimed at improving medication adherence of FH patients.

Participants

Twelve health care professionals from 6 different Dutch hospitals participated in the study. These professionals were recruited by means of a snowball sampling. All participants actively treated people with FH. The study included 6 internists, 2 internists in training, 1 rheumatologist in training, and 2 nurse practitioners. Their clinical experience ranged from 2 months to more than 10 years. Eight participants were female, and 4 were male.

Materials

MIK was created through an iterative HCD approach [26,27], involving FH patients and health care professionals throughout the design process to ensure that the design met the needs of both user groups. A prototype of MIK was built with Invision. The advantage of this mock-up way of prototyping was that it allowed quick evaluation before putting efforts in developing the actual software. Hence, the prototype was not fully functional, but it offered an appropriate level of interactivity to have the participants experience the envisioned functionality. The prototype was built to be compatible with a computer screen-size resolution of 1920 x 993 pixels.

The prototype consisted of 4 sections and an overview page:

- Patient profile, that is, details about the patient’s demographics such as name, age, gender, and address, as well as basic medical information such as the diagnosis, medical history, and family anamnesis (Figure 2).
- Measurements, that is, health measurements conducted, reported, and managed by patients themselves over time. These measurements were meant to trigger patients’ necessity beliefs and concerns about side effects to be discussed in the consultation, as well as to directly foster management of their disease. This measurements section consisted of 2 main functions (Figures 3 and 4):
  - Self-reported patient information on experienced side effects, quality of life (ie, EQ-5D [28]), self-reported medication adherence, and previous medication decisions.
  - A visual overview of clinical measurements, including cholesterol levels, blood pressure (BP), and body mass index (BMI).
- Patient preferences, that is preferences of patients regarding treatment, with the aim to make patients more aware of their options and to foster a discussion about their (necessity) beliefs and concerns, and their own preferences (SDM-like approach). This section consisted of 3 main functions (Figures 5 and 6):
  - (List of) topic(s) the patient wants to discuss during the consultation: Patients are required to create a list of a top 3 topics.
  - Treatment preferences of the patient: Patients are required to create a list of the top 3 of their treatment preferences (ie, taking medication, weight loss, smoking cessation, etc); thereby prioritize the options they believe are feasible to reduce their risk of CVDs.
  - Overview of all medication options: The physician can use this feature of MIK during the consultation as a support tool to explain the dosage equivalence of different types of statins and the risk-reducing effects of medication versus lifestyle changes.
- Task list, that is a list of tasks to be agreed upon by the patient and the health care professional. This page could be used to create a list of tasks for the patient for their next consultation, decided on together with the health professional during the consultation. The list show which tasks have been completed and which ones have yet to be completed (Figure 7).
- Overview page, that is one page with the most important information at a glance. This page includes the patient profile, visual overview of the measurements, treatment preferences, and the 2 tasks that are on top of the task list (Figure 8).
Figure 2. Screen patient profile.

Figure 3. Screen measurements.
Figure 4. Screen cholesterol level.

Figure 5. Screen treatment preferences.
Figure 6. Screen medication options.

![Medication Options Screen](image)

Figure 7. Screen task list.

<table>
<thead>
<tr>
<th>Taken</th>
<th>Date</th>
<th>Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 November 2019</td>
<td>☐ Inschrijven bij sportclub (doelsetvlakken)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>☐ Navigeren familiegeschiedenis, oorzaak van overlijden oma janitor</td>
</tr>
<tr>
<td></td>
<td>6 March 2019</td>
<td>☑ Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod</td>
</tr>
<tr>
<td></td>
<td></td>
<td>☑ Lorem ipsum dolor sit amet, consectetur adipiscing elit</td>
</tr>
<tr>
<td></td>
<td>7 December 2019</td>
<td>☑ Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod</td>
</tr>
<tr>
<td></td>
<td></td>
<td>☑ Lorem ipsum dolor sit amet, consectetur adipiscing elit</td>
</tr>
<tr>
<td></td>
<td>3 July 2018</td>
<td>☑ Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod</td>
</tr>
<tr>
<td></td>
<td></td>
<td>☑ Lorem ipsum dolor sit amet, consectetur adipiscing elit</td>
</tr>
</tbody>
</table>
Procedure

Information about one fictional patient case was entered and presented in the prototype of MIK. This fictional case was created based on previous interviews and observations with patients in a pilot study to ensure credibility. The evaluations existed of individual sessions with health professionals, combining a user test with a semistructured interview. The user test allowed an open approach in which the participants were triggered to provide their own perspective. The researcher started with a brief introduction of the aim and context of the study. The participant was then invited to explore the overview page of the prototype. Next, the participant was provided with several task scenarios and invited to play-act these scenarios using the prototype. The researcher took the role of the patient in these playacts. The evaluation was concluded with a semistructured interview, addressing topics such as communication with and relation to the patient, information needs of the professional, implementation and integration with hospital software, and time management (Multimedia Appendix 1). Near the end of the interview, the participant was provided with a sheet that displayed the 8 different functions of MIK, and he or she was asked to rank their top 3 most valuable functions (Multimedia Appendix 2). A different weight was assigned to the first, second, and third most important function, as assigned by the participant, after which the sum of the weighted scores was calculated. This could help prioritize the functions that the designer could focus on and facilitate an objective discussion within the project team. The session took place at the hospital where the professional practiced. The duration of the sessions varied between 25 min and 69 min, with an average duration of 43 min.

Analysis

Each session was audio-recorded and transcribed verbatim by the first author. A qualitative thematic analysis was performed on the data through the process of coding in 6 phases to create established, meaningful patterns: familiarization with data, generating initial codes, searching for themes among the codes, reviewing themes, defining and naming themes, and producing the final report [23]. Data were coded using the Saturate app, a Web-based tool for collaborative qualitative analysis. During the initial open coding, a total of 297 codes was generated. This large number of initial codes can be attributed to the variety of topics discussed during our semistructured interviews and the level of detail in our coding process. Consequently, axial coding was used to aggregate the codes into preliminary themes. For these preliminary themes, we used the 8 functionalities of the prototype (meaning each initial code was transferred to at least one functionality or discarded).

Within each preliminary theme, we separated the codes based on whether it was a positive statement regarding the functionality or rather a statement suggesting a point of improvement regarding the functionality. Next, we decided to look for overarching themes between the functionalities that related to the design and impact of our app to provide insights that are useful for other developers in the future. This resulted in 9 subthemes relating to the perceived effect of our design (eg, making experienced complaints and side effects tangible and negotiable), the appearance of our design (eg, visualizations of clinical results over time), and information provided by our design (eg, an
indicating treatment preferences). To increase the reliability of the coding process, triangulation was used. Three consensus meetings were held with 3 coauthors (CB, OD, and MM) to discuss the codes and themes. They all read 3 interviews, of which one interview was the same for each coder, to look for information in the transcripts that might be contradictory to the described themes. On the basis of these meetings, we eventually agreed upon aggregating the 9 subthemes into 3 final themes as presented below.

Results

Health Care Professionals’ Assessment of the Functionalities Aimed at Improving Medication Adherence

Table 1 shows the participants’ assessments of the different (sub)functionalities provided by our prototype of MIK. One of the top 3 most valuable functions, as indicated by 9 out of 12 participants, was having an overview of what the patient wants to discuss during the consultation, followed by having information on side effects and quality of life as experienced by the patient (8 out of 12 participants).

Perceived Impact of MIK on Patient-Physician Collaboration

Indication of Topics the Patient Prefers to Discuss

The fact that MIK offers the opportunity to the patients to highlight topics the patient wants to discuss during the consultation was considered a good starting point for the consultation by the participants. The participants argued that knowing a patient’s request for help was highly important to provide optimal patient care:

This is in principle the patient’s request for help at that moment in time. Therefore, I believe that is the most important part. [HCP6]

Gaining Insight Into the Treatment Preferences of the Patient

According to the participants, insight into the patients’ preferences regarding treatment could be used to assess whether maladaptive beliefs or misconceptions exist concerning the different treatment options. If this would be the case, the health care professional could provide patient-specific information to correct these misconceptions or beliefs. In addition, the participants also thought they could help motivate the patient to achieve a certain goal when being aware of the patient preference (ie, weight loss). Knowledge concerning the patient’s treatment preferences was also considered to be of value, as there could be a discrepancy between what the patient and what the health care professional prefers regarding treatment. Participants reasoned that information would support them in forming and delivering suitable treatment advice:

Sometimes I can really be taken by surprise. I have my statins ready and the patient says, no way I am not going to take those. Then you start deliberating about how am I going to bring this across well. [HCP3]

I’d like to know beforehand. We can be confronted during consultations with yes, that and this will not work and then you must improvise about what (medication) to give. [HCP11]

Making Experienced Complaints and Side Effects Tangible and Negotiable

The feature in MIK which can collect and display information about the patients’ experienced side effects and quality of life was considered important by the participants as they know from experience that a poor quality of life or (unacceptable) side effects could interfere with adherence and, thereby, the positive effects of the treatment for FH. During the prototype test, more than half of the participants (7/12) noticed that the fictional patient presented in MIK experienced side effects, which led them to inquire about this during the consultation. According to the participants, it could prompt a more open discussion about patient’s daily functioning (Figures 2 and 8). They imagined that patients would feel less burdened and hesitant in bringing forward their complaints using the app compared with doing this face-to-face in the consultation. Moreover, addressing these issues with the patient was seen as a necessity to prevent nonadherence.

Addressing certain complaints or questions will be improved dramatically. [HCP10]

Well, I believe it is a good thing that people arrive at the consultation prepared. And things are addressed in this manner. There are also people who do not give notice (about complaints). Or they do not dare. Or they are ashamed about it. [HCP10]

Creating a Task List Together

Setting tasks together during the consultation was seen as a feature which could stimulate the patient-physician collaboration and patients’ self-management skills. A list of tasks available and visible in the consultation was considered to be a clear way to see what tasks are still pending and what tasks were already completed. It was suggested that the ability to access the task list at any place and any time could provide patients with more control while simultaneously serving as a reminder for the next consultation.

Perceived Impact of MIK on Patient’s Understanding and Self-Management

Active Role for the Patient in Collecting and Providing Information

In the MIK user scenario, the patients prepared themselves at home before the consultation by filling in the information on side effects, quality of life, and treatment preferences. Participants believed this could lead to patients taking a more active role in their treatment by increasing their self-awareness regarding their condition and treatment regimen. They also argued that this would make it easier for the health care professionals in helping and encouraging patients to reach their goals (ie, losing weight).
Table 1. Number of participants (n=12) who qualified the different prototype functions as the first, the second, or the third most valuable in their daily practice. Between brackets the summation of the weigh factors (wf) is depicted: the most valued function receives a weighing factor of 3, the second most valued function receives a weighing factor of 2, and the third most valued function receives a weighing factor of 1.

<table>
<thead>
<tr>
<th>Functionalities</th>
<th>Most valued function, n (wf)</th>
<th>Second most valued function, n (wf)</th>
<th>Third most valued function, n (wf)</th>
<th>Total, n (wf)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview page</td>
<td>2 (6)</td>
<td>2 (4)</td>
<td>1 (1)</td>
<td>4 (10)</td>
</tr>
<tr>
<td>Patient profile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurements</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detailed qualitative patient information (side effects, quality of life)</td>
<td>4 (12)</td>
<td>2 (4)</td>
<td>2 (2)</td>
<td>8 (18)</td>
</tr>
<tr>
<td>Visual overview of cholesterol, BP(^a), and BMI(^b)</td>
<td>3 (9)</td>
<td>1 (2)</td>
<td></td>
<td>4 (11)</td>
</tr>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overview of topic(s) the patient wants to discuss</td>
<td>3 (9)</td>
<td>4 (8)</td>
<td>2 (2)</td>
<td>9 (19)</td>
</tr>
<tr>
<td>Treatment preferences</td>
<td></td>
<td>2 (4)</td>
<td>5 (5)</td>
<td>7 (9)</td>
</tr>
<tr>
<td>Overview of all medication options</td>
<td></td>
<td>1 (1)</td>
<td></td>
<td>1 (1)</td>
</tr>
<tr>
<td>Task list</td>
<td>1 (2)</td>
<td>1 (1)</td>
<td></td>
<td>2 (3)</td>
</tr>
</tbody>
</table>

\(^a\)BP: blood pressure.
\(^b\)BMI: body mass index.

**Visualizations of Clinical Results Over Time**

Participants described the visual graphs of the changes in body weight, cholesterol, and BP over time as simple and easy in appearance and interpretation. They reasoned that the visual graphs, especially in combination with the verbal information provided during the consultation, would be useful in objectifying the results of the treatment in terms of risk-reducing effects, thereby potentially enhancing a patients’ understanding of their disease. This was considered important, as FH patients usually do not notice changes in their BP or cholesterol in day-to-day life. Interestingly, the professionals showed different ways of using the graphs during the playact consultation. The graphs were used to encourage the patient to keep up the good work (ie, complying with medication) and to make the patient more aware of the risks associated with their current weight, BP, and cholesterol levels (ie, the high risk of CVD or diabetes when keeping this weight). Participants also acknowledged the importance of including the target cholesterol level that should be achieved as per the current clinical guidelines. High, moderately high, and normal cholesterol levels were displayed using “traffic light colors,” which were considered to be an essential piece of knowledge for the patient.

*Then people can review and look back, if I deliver effort, it will be rewarded in the numbers. And that can be really motivating.* [HCP6]

*And if you take those (pills), you clearly see it is decreasing. And here you see a value in the green area. And the green area is the guideline. Because if the LDL-cholesterol is beneath the 2.5, that is really the goal of the treatment. It is the Dutch protocol and you can reach it (points at circle in green zone). If you use your medication and mind you diet.* [HCP2]

**Perceived Impact on Clinical Decisions and Workflow**

Participants expressed ambivalent thoughts on how MIK (ie, its general usage and its different functionalities) could contribute to an effective and efficient workflow and consultation.

**Interpreting Measurements in a Bigger Context and Over Time**

Participants valued having insight into the various measurements over time (Figure 2). They argued that it could help them quickly identify connections between the measurements and other patient data (ie, quality of life score), which could optimize their clinical decision making (eg, changing the type of statins used or referring patients to a dietician or psychologist for support in losing weight or how to cope with their condition).

*We are always looking for patterns and links, and there has been a change in medication and I see some colours have changed. This could mean there is a causal relation.*

In addition, participants felt supported by having all patient information together in one overview (the measurements overview page presented information about the patient's medication history, experienced side effects, quality of life, medication adherence, smoking, BMI, cholesterol, and BP). They argued that this could help them prioritize what topics need their attention before the start of the consultation.

*I believe that, look if someone is not feeling well, it can just be difficult to stick to therapy. So in this case I would definitively ask and see how, well, what we can do about it and if we should directly act on it.* [HCP9]

*I am here to improve someone's health, but something that bothers the patient enormously can also be in the way of the medical treatment. So that is something*
that you should be able to address (about quality of life). [HCP5]

More Effective and Tailored Consultation

Participants believed that MIK could make the 15-min consultation with FH patients more effective, as they would be able to spend less time asking the patient standard questions about the patients’ well-being. With MIK, patients would have already answered those questions and, therefore, important information could be reviewed by the health professional before the consultation. Together with the overview of all measurements, participants believed this would positively contribute to an effective and efficient preparation of the consultation. Another positive aspect was that the health care professionals would be more aware of the topics the patient wants to discuss. This could be time-saving and, thus, be an incentive for the health care professional to use the app.

I think that if I know what the patient wants to discuss, if can save me time. Sometimes it can take a while before the word is out. And now it can be much more efficient, if we know immediately what we want to address. That would be a reason for me to look beforehand for 2 seconds like...are there any highlights that need to be addressed. [HCP6]

Additional Workload

Participants described various undesirable aspects of using an app that runs separately from the electronic patient record system. A major negative aspect was having to work with 2 systems, the electronic patient record system and MIK. They did not prefer a situation in which extra actions would be required, such as logging in and finding the right patient in MIK.

Two systems, that I would find a big disadvantage. You notice this now with many apps, all need extra actions, so that would be the biggest drawback. [HCP11]

Even though most of the work of the extra registration would be on behalf of the patient, some participants disliked the fact that there would be a double registration of the laboratory values, and they would have to learn to work with a new program. The use of 2 screens in a consultation was also seen as distracting by some health care professionals, and information communication technology (ICT) prerequisites (ie, results not coming through) were regarded as something to be avoided. Besides the potential extra system and the potential distraction of having an extra screen, one participant feared that patients would expect their physician to read and act upon all the information supplied by MIK, despite the limited time available. Particularly, concerning the topic of quality of life, several participants were doubtful on how to deal with this “broad” information and considered the quality of life-related issues beyond the scope of consultation with a vascular specialist.

Something you surely want to avoid is the patient to pour out their heart in those 10 minutes of the consultation. That is something I am doubtful about. [HCP8]

Large pieces of textual information submitted by the patient in the various comment boxes were considered inefficient, as this would be time-consuming to read. Additionally, one participant remarked that patients may have difficulty expressing their thoughts in writing. Another participant explained that textual information in the app might be more difficult to assess in terms of importance and severity compared with a face-to-face story. Another perceived disadvantage concerned the fact that when data from MIK would be exported and saved in the hospital system, modifications would no longer possible. In this respect, the interactivity of the app would be lost, which was argued to have a negative influence on the workflow.

Reliability of Data Provided by Patients

There were differences in opinions regarding the reliability of the self-reported patient data in MIK. On the one hand, health care professionals argued that registration of laboratory results was less reliable when done by the patient, while on the other hand, professionals suggested that patients would be more accurate when given the opportunity to keep track of their own data.

Discussion

Principal Findings

This explorative qualitative study investigated the perspective of health care professionals concerning a newly developed digital app—MIK—aimed at improving the medication adherence of patients with FH. By means of the 4 functionalities in the app (ie, patient profile; health measurements [experienced side effects and quality of life]; treatment preferences; and task list), MIK was targeted at improving patient engagement, self-management, and SDM in consultation. This study showed that these targets are largely feasible, based on the perspective of professionals involved in the care for FH patients. The majority of professionals argued that the app could improve the focus and efficiency of the consultation, enhance patient engagement, and even influence the treatment decisions made, indicating a potential shift toward SDM.

Although most of the participants were positive regarding the functions of MIK, there were considerable differences between health professionals in the degree and the manner in which the information in MIK would be used in the consultation. Some professionals argued that they would use the information in MIK to engage with the patient and prompt discussions about the patients’ beliefs and concerns so as to detect or decrease intentional nonadherence. Other professionals were more likely to use the information in MIK to provide more personalized clinical treatment. While the underlying reasons behind this difference deserve further evaluation, there is not necessarily a right or wrong way of using the information in MIK. We would like to stress that SDM is a continuum rather than a fixed way of sharing decisions with patients, and it will necessarily take different forms in different decision situations [29]. In this case of FH, there is an expert agreement that taking lipid-lowering medication, especially statins, is superior to any other treatment option (ie, lifestyle changes or homeopathic products). Hence, taking statins for the treatment of FH is not considered to be a strict-preference-sensitive decision. This is different compared...
with situations where multiple eligible options exist from a medical perspective that involve a trade-off among different possible outcomes of each treatment [16]. However, SDM can also be an appropriate model for FH patients because the options that may be nonequivalent as per the medical experts do seem to be sensitive to the patients’ own preferences. Moreover, within the statin regimen, there are also multiple decisions to be made with respect to the type or dosage, which should obviously be discussed between patients and professionals to target intentional nonadherence. The aim of MIK is to stimulate a more open discussion about the patients’ beliefs and preferences and to take these beliefs and preferences more explicitly into account when deciding on the type and dosage of statins and lifestyle changes. Generally, the interviewed professionals indicated that they would indeed use the information in MIK to prevent or target intentional nonadherence.

Health professionals particularly embraced the information about patients’ experienced side effects and quality of life, as well as the information about patients’ treatment preferences. This information is not routinely discussed in the current care process with FH patients. According to the interviewed professionals, this information can prompt a discussion about patients’ beliefs and concerns and can correct misconceptions and fill knowledge gaps regarding the different treatment options. The positive attitude among professionals toward patient-reported outcome (PRO) measures, such as quality of life, also seems to fit the current perspectives in health care that PROs can be equally important as clinical outcomes (ie, value-based health care) [30]. Information on PROs cannot be extracted from patients’ medical records or a proxy; therefore, PROs need to be assessed in their own right [31]. The PROs may provide important additional information to professionals to reach a more individualized patient approach. Our interviewed professionals especially addressed the interconnection between patients’ quality of life and anthropometric levels (ie, cholesterol, BP, and weight) as this could indicate low (intentional) adherence.

This study showed that user testing with health professionals resulted in valuable design implications. For example, the professionals stressed the importance of (audio–visual options for explaining different types of statin medication to increase patient understanding. Professionals also made specific suggestions (which were not all described in detail above), for example, about the use of specific colors and shapes to make the app more intuitive, about the presentation of a cutoff level of cholesterol in the graph and about navigation between the screens of the app. These suggestions can be used to optimize the usability of the app. However, other design implications were more focused on the integration of the app with the electronic patient record system to avoid the use of 2 screens and the need for double data entry. Unfortunately, this problem is encountered by many studies focused on eHealth innovations and not easily resolved [32,33].

**Limitations and Further Recommendations**

Although our participants were practicing health professionals, we were not able to evaluate the digital app in a real consultation with patients and their professionals. Instead, we asked professionals to imagine themselves being in the hypothetical situation that they were preparing for a consultation with a fictitious patient. It is possible that in actual practice, other issues will emerge that have not been captured in our study. In addition, the fact that the tool was a Web-based mock-up rather than a fully functioning app had its limitation. While this click-through mock-up was able to show the interface and the most important features of the app, some of the option buttons were disabled. Furthermore, although the interviewed professionals worked in 6 different academic and top clinical teaching hospitals, we cannot assume their perspective is representative for all professionals treating FH patients in the Netherlands. More research is warranted to evaluate how MIK supports a larger group of professionals in practice. In addition, we strongly recommend the need to evaluate the effect of digital tools on patient outcomes such as medication adherence, satisfaction with care, health status, and morbidity. In the past few years, there has been an enormous expansion of digital tools, for example, mobile apps, to support patients in taking their medication. Recent evaluations have shown that although some of those apps are of good quality, the effectiveness of these tools with regard to patient outcomes, such as adherence, remains unknown [34]. Finally, we recommend the development of general user-centered design principles for developing eHealth apps to optimize medication adherence. These design principles allow research institutes and design agencies to design eHealth apps for patient engagement, self-management, and medication adherence or to enhance the applicability and usability of their eHealth tools for other health disorders.

**Conclusions**

The interviewed professionals largely embraced MIK arguing that the app could improve the focus and efficiency of the consultation and even influence treatment decisions made. They particularly valued the information about patients’ experiences with side effects and about their quality of life, which is information that is not routinely discussed in the current care process but could prompt a discussion about patients’ beliefs and concerns. According to the professionals, MIK can be used to discuss the options that exist within a treatment regimen more explicitly. Professionals also acknowledged the self-management function of MIK, making connections between data would engage and motivate patients outside the consultation to adhere to their treatment.

**Acknowledgments**

This work was financed by Design United, the Dutch research center for design (Project C2D1604).
None declared.

Multimedia Appendix 1
Interview guide.

[PDF File (Adobe PDF File), 31KB - humanfactors_v5i1e7_app1.pdf]

Multimedia Appendix 2
Function sheet.

[PDF File (Adobe PDF File), 21KB - humanfactors_v5i1e7_app2.pdf]

References

Abbreviations

BMI: body mass index  
BP: blood pressure  
CVD: cardiovascular disease  
eHealth: electronic health  
FH: familial hypercholesterolemia  
HCD: human-centered design  
ICT: information communication technology  
NCF: Necessity–Concerns Framework  
PROs: patient-reported outcome  
SDM: shared decision making
Evaluating the Usability and Usefulness of a Mobile App for Atrial Fibrillation Using Qualitative Methods: Exploratory Pilot Study

Jaclyn Hirschey1*, BA; Sunetra Bane1*, MPH; Moussa Mansour2, MD; Jodi Sperber1,3, PhD; Stephen Agboola1,3,4, MPH, MD; Joseph Kvedar1,3,4, MD; Kamal Jethwani1,3,4, MPH, MD

1Partners Connected Health, Partners Healthcare, Boston, MA, United States
2Heart Center, Massachusetts General Hospital, Boston, MA, United States
3Harvard Medical School, Cambridge, MA, United States
4Massachusetts General Hospital, Boston, MA, United States
*these authors contributed equally

Corresponding Author:
Jaclyn Hirschey, BA
Partners Connected Health
Partners Healthcare
25 New Chardon Street
Suite 300
Boston, MA, 02114
United States
Phone: 1 617 724 2724
Email: jhirschey@mgh.harvard.edu

Abstract

Background: Atrial fibrillation (AFib) is the most common form of heart arrhythmia and a potent risk factor for stroke. Nonvitamin K antagonist oral anticoagulants (NOACs) are routinely prescribed to manage AFib stroke risk; however, nonadherence to treatment is a concern. Additional tools that support self-care and medication adherence may benefit patients with AFib.

Objective: The aim of this study was to evaluate the perceived usability and usefulness of a mobile app designed to support self-care and treatment adherence for AFib patients who are prescribed NOACs.

Methods: A mobile app to support AFib patients was previously developed based on early stage interview and usability test data from clinicians and patients. An exploratory pilot study consisting of naturalistic app use, surveys, and semistructured interviews was then conducted to examine patients’ perceptions and everyday use of the app.

Results: A total of 12 individuals with an existing diagnosis of nonvalvular AFib completed the 4-week study. The average age of participants was 59 years. All participants somewhat or strongly agreed that the app was easy to use, and 92% (11/12) reported being satisfied or very satisfied with the app. Participant feedback identified changes that may improve app usability and usefulness for patients with AFib. Areas of usability improvement were organized by three themes: app navigation, clarity of app instructions and design intent, and software bugs. Perceptions of app usefulness were grouped by three key variables: core needs of the patient segment, patient workflow while managing AFib, and the app’s ability to support the patient’s evolving needs.

Conclusions: The results of this study suggest that mobile tools that target self-care and treatment adherence may be helpful to AFib patients, particularly those who are newly diagnosed. Additionally, participant feedback provided insight into the varied needs and health experiences of AFib patients, which may improve the design and targeting of the intervention. Pilot studies that qualitatively examine patient perceptions of usability and usefulness are a valuable and often underutilized method for assessing the real-world acceptability of an intervention. Additional research evaluating the AFib Connect mobile app over a longer period, and including a larger, more diverse sample of AFib patients, will be helpful for understanding whether the app is perceived more broadly to be useful and effective in supporting patient self-care and medication adherence.

(JMIR Hum Factors 2018;5(1):e13) doi:10.2196/humanfactors.8004

KEYWORDS
nonvalvular atrial fibrillation; medication adherence; patient self-care; mobile application; exploratory research; pilot study; usability study; acceptability study; qualitative methods

http://humanfactors.jmir.org/2018/1/e13/
Introduction

Atrial fibrillation (AFib) is the most common type of heart arrhythmia [1]. It is estimated that in the United States between 2.7 to 6.1 million people currently have AFib and that 1 in 4 adults 40 years and older will develop AFib during their lifetime [1,2]. It is characterized by palpitations, dizziness, weakness, and dyspnea and associated with increased health care costs and mortality and reduced quality of life [3,4]. Additionally, individuals with AFib have a 4- to 5-fold increased risk of stroke [5].

To manage AFib stroke risk, more than half of all individuals in the United States with AFib are prescribed a nonvitamin K antagonist oral anticoagulant (NOAC) [6]. NOACs have several advantages over older vitamin K antagonist anticoagulant medications, such as warfarin, because of their lower risk for food and drug interactions, simpler dosing regimens, and lack of requirement for continuous blood monitoring [7]. However, medication nonadherence—a common issue among many chronic conditions—continues to be a challenge for NOAC treatment [8,9]. More than half of individuals on an NOAC for AFib do not meet the Pharmacy Quality Alliance adherence threshold of 80%, putting them at an increased risk for thrombus formation [9,10]. Additionally, while underanticoagulation may pose a greater risk for stroke, overanticoagulation can increase the risk of bleeding [11]. Thus, careful adherence to clinician-prescribed treatment is essential to keep within a therapeutic dosing range and prevent adverse events.

The need for strict treatment adherence, coupled with distressing symptoms and disease complexity, make patient self-care difficult [12]. Although the introduction of NOACs has reduced the patient burden associated with warfarin treatment, it has also highlighted the need for new tools that support self-care and treatment adherence in the absence of frequent clinical oversight [13,14]. Existing tools to support AFib anticoagulant treatment have largely focused on providing decision support to clinicians at the point of prescription [15-17]. Additional patient-facing tools that target medication adherence and long-term self-care may be valuable, particularly for patients taking NOACs [13].

The AFib Connect mobile app (Figure 1), created for both Android and iPhone operating system (iOS) platforms, was developed with the goal of supporting long-term patient self-care and adherence to anticoagulant therapy. The app was developed by an interdisciplinary design team of clinicians, qualitative researchers, and user experience designers at Partners Connected Health in collaboration with Daiichi Sankyo, Inc. As part of a user-centered design approach, input from clinicians and patients was compiled to understand the primary goals, needs, and preferences for the app [18]. Semistructured interviews were conducted with nine AFib clinicians and patients to identify the app’s core features. An iterative process of feedback from key stakeholders was used to refine the app’s overall design. Usability testing was then conducted with clinicians and patients in a lab using the first version of the app. This feedback was incorporated back into the design of version two, which was used for this study. Table 1 outlines the features included in version two of the AFib Connect mobile app.

The goal of this pilot study was to evaluate the perceived usability and usefulness of the AFib Connect mobile app after an extended period of natural use by AFib patients prescribed an NOAC.
**Figure 1.** Dashboard screen of the AFib Connect app.

**Table 1.** AFib Connect app feature list and descriptions.

<table>
<thead>
<tr>
<th>Features</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFib(^a) Guide</td>
<td>An introduction to AFib through text and animated videos, including an overview of the condition, associated stroke risk, and a decision tool to review treatment options; information provided in the guide is based on the American Heart Association and Massachusetts General Hospital guidelines</td>
</tr>
<tr>
<td>Library</td>
<td>Detailed information on AFib, including types of medication available, procedure options, and guidance on medication adherence and stroke risk; information provided in the Library is based on the American Heart Association and Massachusetts General Hospital guidelines</td>
</tr>
<tr>
<td>Episode Tracker</td>
<td>Patient-generated log for tracking AFib episodes and associated notes for documentation and review with physician</td>
</tr>
<tr>
<td>Trigger Tracker</td>
<td>Patient-generated log for tracking possible episode triggers (eg, caffeine, alcohol, and poor sleep)</td>
</tr>
<tr>
<td>News Feed</td>
<td>Curated news content from five heart health–related Twitter feeds, such as American Heart News, StopAF.org, and the American Heart Association</td>
</tr>
<tr>
<td>Medication Reminder</td>
<td>Reminders to take medication at a designated time, including pop-up notifications, an option to mark medications as taken, and adherence history</td>
</tr>
<tr>
<td>Heart Rate Monitor</td>
<td>Tool for measuring heart rate using the mobile phone camera</td>
</tr>
<tr>
<td>Appointment Reminder</td>
<td>Calendar for tracking medical appointments and reminders of upcoming visits</td>
</tr>
</tbody>
</table>

\(^a\)AFib: atrial fibrillation.
Methods

This pilot study used qualitative methods and an exploratory research approach that combined naturalistic app use, surveys, and semistructured interviews to understand patient perceptions of the mobile app. A 4-week, five-visit study design gave participants the opportunity to use the app in their everyday environment and provide detailed feedback on select app features each week. The study was conducted from September 2016 to April 2017. Participants were asked background questions at weekly intervals over the 4-week study period. A total of five interviews were conducted with each participant.

Study Population

The study was conducted from September 2016 to April 2017. A purposeful sample of individuals diagnosed with nonvalvular AFib and taking NOACs to manage stroke risk were selected to participate. Potential participants were identified by clinician referral from the department of cardiology at Massachusetts General Hospital and contacted for recruitment. A total of 16 participants enrolled and 12 participants completed the 4-week study. Among the 4 participants who did not continue through study closeout, 2 were found ineligible after enrollment because of mobile phone operating system incompatibility, 1 was lost to follow-up after week 1, and 1 dropped out after enrollment because of a lack of interest. Enrollment of participants continued until thematic data saturation was reached.

Data Collection

Surveys were administered in-person at study enrollment and by mail at closeout. Semistructured interviews were conducted in-person during the enrollment visit to establish rapport between the participant and the interviewers and by phone for the remainder of the study for participant convenience. Interview data collected from participants was deidentified before data storage and analysis. The AFib Connect app was downloaded by research staff onto each participant’s personal mobile phone for use throughout the study period. Any protected health information that a participant may have entered into the app was securely stored locally on their phone and was not accessible to the researchers. Table 2 provides an overview of the study design and data collection schedule.

Semistructured Interviews

A total of five interviews were conducted with each participant at weekly intervals over the 4-week study period. During enrollment, participants were asked background questions regarding their AFib history, overall technology use, and expectations about using an app for AFib. Participants were then asked to explore the app on their mobile phones and provide their initial impressions of each feature and the app overall. Observations about the participant’s interaction with the app were noted by the researchers.

From week 1 to 3, participants were asked to explore 2 or 3 predetermined features in detail in addition to the app overall. A list of these features and brief instructions were emailed to participants 1 week before the interview as a reminder. Phone interviews lasted 30 min, and participants were asked to discuss and rate the features they tested over the past week and the app overall in terms of its usability and current usefulness to them. During week 4, participants were asked to again review the entire app and provide feedback and a rating on their overall experience. Throughout the interviews, participants were encouraged to provide their honest and candid feedback about the app. Researchers paid close attention to conversational tone and pauses and asked follow-up questions, where needed, to probe more deeply into participant’s responses and to minimize any respondent bias.

Each interview was attended by two qualitative researchers, with one researcher leading the interview and the other taking detailed notes. Interviews were audiorecorded, and transcriptions of each recording were generated. At the completion of every interview, notes were discussed and summarized by the researchers. Utilizing grounded theory, a coding framework was developed from the interview questions. At regular intervals throughout the study, emergent codes were derived from note summaries and interview transcripts. After study completion, codes from each interview were compared and then organized into themes to derive the final results.

Surveys

Supplementary study data was collected by custom surveys previously developed by Partners Connected Health. At study enrollment, information on participant demographics and technology use was collected. At closeout, patient satisfaction and app usability was measured with a 5-point Likert scale, yes or no, and open-ended questions. Survey data was analyzed for the 12 participants who completed the study. Demographic, technology use, patient satisfaction, and usability characteristics were summarized. Descriptive statistics were reported as means and standard deviations for continuous variables and as percentages for categorical variables.
Table 2. Study design and data collection schedule.

<table>
<thead>
<tr>
<th>Study visit and interviews</th>
<th>Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Enrollment (in person)</strong></td>
<td></td>
</tr>
<tr>
<td>Medical and AFib(^a) history</td>
<td>Demographics</td>
</tr>
<tr>
<td>Current use of technology</td>
<td>Technology use</td>
</tr>
<tr>
<td>Expectations from app</td>
<td></td>
</tr>
<tr>
<td>Initial impression of each feature and app overall</td>
<td></td>
</tr>
<tr>
<td><strong>Week 1(^b) (by phone)</strong></td>
<td>N/A(^c)</td>
</tr>
<tr>
<td>AFib Guide</td>
<td></td>
</tr>
<tr>
<td>Medication Reminder</td>
<td></td>
</tr>
<tr>
<td>Additional feedback by feature and app overall</td>
<td></td>
</tr>
<tr>
<td><strong>Week 2(^b) (by phone)</strong></td>
<td>N/A</td>
</tr>
<tr>
<td>Heart Rate Monitor</td>
<td></td>
</tr>
<tr>
<td>Episode Tracker</td>
<td></td>
</tr>
<tr>
<td>Library</td>
<td></td>
</tr>
<tr>
<td>Additional feedback by feature and app overall</td>
<td></td>
</tr>
<tr>
<td><strong>Week 3(^b) (by phone)</strong></td>
<td>N/A</td>
</tr>
<tr>
<td>Trigger Tracker</td>
<td></td>
</tr>
<tr>
<td>News Feed</td>
<td></td>
</tr>
<tr>
<td>Appointment Reminder</td>
<td></td>
</tr>
<tr>
<td>App overall</td>
<td></td>
</tr>
<tr>
<td><strong>Closeout (by phone)</strong></td>
<td></td>
</tr>
<tr>
<td>Needs when first diagnosed</td>
<td>Satisfaction and usability</td>
</tr>
<tr>
<td>Likelihood to use after study</td>
<td></td>
</tr>
<tr>
<td>Likelihood to recommend to others with AFib</td>
<td></td>
</tr>
<tr>
<td>Additional feedback by feature and app overall</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)AFib: atrial fibrillation.
\(^b\)Detailed feedback on the features schedueld for data collection.
\(^c\)N/A: not applicable.

**Results**

**Participant Characteristics**

The study comprised 7 males and 5 females, ranging in age from 37 to 67 years, with a mean of 59 years. Participants had been managing their AFib for 6 years on average, with a range from 1 to 15 years, and none were newly diagnosed. Eleven out of 12 participants (92%) were asymptomatic at the time of study participation because of having an ablation or cardioversion procedure, or a diagnosis of persistent AFib. One participant (8%, 1/12) experienced an AFib episode during the study. Table 3 provides a summary of characteristics for the study participants.

**App Usability**

Results from survey data showed that all 12 (100%, 12/12) participants somewhat or strongly agreed that the app was easy to use and navigate, with 9 (75%, 9/12) stating they always knew what to do in the AFib app, and only one (8%, 1/12) reported needing to ask for help while using the app. Ten participants (83%, 10/12) somewhat or strongly agreed that the AFib app acted and felt like other apps they had used before.

Interview data revealed a positive perception of app usability; however, participants identified a few areas that could be improved to provide a better overall user experience. Areas of usability improvement can be organized into 3 categories: navigation, clarity of instructions and design intent, and software bugs.

**Navigation**

Participants stated that finding key features and the navigation between screens of the app was simple and straightforward:

*I thought it was well designed as an app in that it sort of follows the typical style of most apps, so you don’t really—it’s easy just to touch things and you understand quickly what you need to be doing...*
thought it was very well designed in terms of navigation. [Participant 5]

However, one participant (8%, 1/12) reported difficulty discovering some of the AFib video content and another could not locate the news feed. There was also a reported disruption of workflow when, after reviewing an article in the Library and returning to the Library home screen, users were brought to the top of the page rather than to the section where they had left off.

Clarity of Instruction and Design Intent

Additional instruction, or clearer design intent in some areas of the app, might also improve the app’s overall usability. During app set-up, the researchers observed that nearly all the participants questioned whether certain data fields were required or optional and the type of information they should to enter in the medication notes field.

Taking a heart rate reading using the mobile phone camera was a novel and liked concept for nearly all the participants; however, many individuals expressed uncertainty about how the feature worked, where to place their finger on the camera flash, and whether they were taking their heart rate correctly. Nine of the 12 (75%) participants mentioned that having step-by-step illustrations of how to use the feature and additional context on how to interpret and act on readings would be helpful:

...It would be helpful if, within the app, there was some information like [normal heart rate range and heart rate range after an ablation] because as I’m taking my heart rate, I’m thinking, “My resting heart rate is supposed to be around 60. Now it’s 80.” So, I had to go outside of the app to get that information. [Participant 7]

The Trigger Tracker was another feature that participants were initially uncertain how to use. Although the feature was designed to allow users to log their triggers upon exposure to establish a trigger history, 7 of the 12 (58%) participants assumed that they would note potential triggers retrospectively only after an episode occurred:

I guess I was a bit confused as to the purpose of this. I read the information a couple of times and I sort of walked away unclear. I mean, I could make assumptions, but I sort of walked away unclear. It seemed like you’re asking me to input triggers so that I can determine what my triggers are. [Participant 15]

Potentially, because of assumptions about intended use or similarities in naming, there was also some confusion about the difference between the Trigger Tracker and the Episode Tracker, with 2 participants (25%, 2/12) referring to them as if they were the same feature.

Table 3. Summary of characteristics for the 12 study participants.

<table>
<thead>
<tr>
<th>Participant characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td>59.25 (7.78)</td>
</tr>
<tr>
<td>Time since AFib diagnosis (years), mean (SD)</td>
<td>5.67 (4.54)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
</tr>
<tr>
<td>12 years or completed high school or general educational development</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Some college</td>
<td>2 (17)</td>
</tr>
<tr>
<td>College graduate</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Graduate or professional degree</td>
<td>6 (50)</td>
</tr>
<tr>
<td>Employment, n (%)</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Homemaker</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Self-employed, full or part-time</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Retired</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Mobile phone type, n (%)</td>
<td></td>
</tr>
<tr>
<td>Android</td>
<td>3 (25)</td>
</tr>
<tr>
<td>iPhone</td>
<td>9 (75)</td>
</tr>
</tbody>
</table>

aAFib: atrial fibrillation.
Software Bugs
Seven out of 12 (58%) of participants reported software bugs that negatively affected their experience with the app. Participants on Android devices reported occasional inconsistent readings or app crashes when using the Heart Rate Monitor. Participants who reported this issue assumed the readings were inaccurate, with 1 participant (8%, 1/12) indicating that seeing this fluctuation caused some minor anxiety. The app error was immediately corrected to prevent further issues, and all the participants were given additional verbal instructions on how to take heart rate readings correctly, as unclear app instructions may have partly contributed to the variation in values.

Two participants (25%, 2/12) also reported a bug on the Appointment Reminder and Medication Reminder calendars, which caused the date drop-down fields to be cut off from view. Although the feature was still usable despite the bug, the affected participants reported a less user-friendly experience because of the issue.

App Usefulness
Overall, 92% (11/12) of participants reported being satisfied or very satisfied with the AFib Connect app, and all the participants indicated that they were somewhat or very likely to recommend the app to people in treatment for AFib. Ten out of 12 (83%) participants somewhat or strongly agreed that they found the AFib app useful as a tool to track AFib related information, manage medical appointments, and be reminded to take their medication:

I’ve had AFib for a year and a lot of the information that I received, [was through my own] efforts of research, and a lot of the tracking information that I needed to monitor my condition was done by hand. [This app] is a convenient way and tool to keep things organized, to handle appointments and reminders, and to interface between yourself, your condition, and your caregivers. [Participant 3]

[By study closeout] I continued to really just utilize the reminder feature, which works well. [It] works just the way I need it to work for me, for my personality, which is it bugs me until I don’t want it to bug me anymore so I take my pills. And frankly, for me, the app’s worth having just for that. [Participant 5]

Although participant perceptions of the app’s usefulness for supporting self-care and medication adherence were largely positive, several areas of improvement were identified. The following outlines participant perceptions of usefulness for each of the app’s eight features and the app overall.

Library
The Library was rated highly in terms of usefulness to participants. Participants reported that the content was clear, relevant, comprehensive:

I think that [there is] a lot of good basic information...it’s easy to understand and informative enough for anyone who has AFib problems. It pretty much covers everything about it. [Participant 14]

Several participants (58%, 7/12) reported that while they were familiar with most of the Library content and might not use it every day, it was still a nice resource for them to reference and refresh their knowledge. Five participants (42%, 5/12) also requested more treatment-related information, including comparative data between newer AFib medications and procedures.

Medication Reminder
Participants also responded positively to the Medication Reminder feature. Ten participants (83%, 10/12) reported that the AFib app helped them keep track of taking their medications. Seven (58%, 7/12) reported using the feature about once a day. More than half of participants reported that even without the AFib app they would have remembered to take their medications every day because of having an existing means or habit of remembering to take their medication:

But I think I just like having a reminder that pops up every day when I’m supposed to take the meds, I think that’s a neat feature...Even though I’d always remember, it’s just nice to have it pop up and remind you. [Participant 10]

One participant (8%, 1/12) identified an additional use for the Medication Reminder feature, indicating that she would use it to keep track of all her medications and dosage information, beyond those she takes for AFib, so that she has this information readily accessible when needed.

Heart Rate Monitor
Despite the usability issues noted previously, participants liked the ability to check their heart rate quickly and easily and keep a history of their readings on their mobile phone. Although 11 out of 12 participants (92%) did not experience episodes during the study period, all of them reported checking their heart rate periodically throughout the day, or after exercise:

...[AFib patients] want to make sure that their heart rate is nice and even and down where it should be. When it’s out of whack, that’s a good indicator that you’re going into AFib. Some people don’t really know they’re in AFib unless they check that. So I think that’s an important part of your application. [Participant 9]

AFib Guide
AFib Guide information about the different types of medication and procedures for AFib would have been especially useful to patients when they were first diagnosed. Participants reported receiving some education about AFib from their doctor; however, most of them indicated that they had to do a lot of their own Internet research, which took some effort. As, on average, our study sample had been living with AFib for 6 years, the Guide was used as more of a reference to review information:

It has a lot of good information in there. A lot of the information I did know already, but some other things that I did not know. I thought it was kind of interesting that you could make a list of what was important to you, and to go over with the doctor...It made me more
aware of the details of AFib, what happens to you. But all of the treatments and the reasons why you have ablation, I was already very familiar with...For me, I would still want to keep it because I just think it’s a good review. [Participant 14]

Participants reported that the AFib Guide’s videos and animations were especially helpful in explaining the content. Additionally, participants liked the idea of emailing their AFib Guide results to a doctor, although they indicated that they would more likely reference this record at an in-person visit.

**Episode Tracker**

As our study participants were asymptomatic at the time of the study, the Episode Tracker feature was not relevant to their current condition. Still, participants liked the idea of tracking an episode’s duration and making notes of what happened and felt it would have been a useful record-keeping tool when they were first diagnosed:

> I would definitely use the episode tracker because that would eliminate my need for writing lengthy notes on my iPhone. So, I thought that the episode tracker was set up very well, and would probably shorten the length of time needed to get the information down where I could access it quickly and possibly reference it or email it to the appropriate party...I’m not using it now because I’m not having episodes. When I was having episodes, I believe I would use it. [Participant 3]

Most participants indicated that if they believed they were having an episode they would first take their heart rate to verify if it’s elevated; if yes, they would then begin recording the episode’s duration and make note of any potential triggers. One participant (8%, 1/12) also suggested that the Episode Tracker include information to help individuals get through an episode, for example, by encouraging them to breathe slowly.

**Trigger Tracker**

Although participants liked the idea of the Trigger Tracker, this feature had less relevance to them during our study as most were not experiencing episodes. Some individuals also suggested that the Trigger Tracker feature would be more useful if it was combined with the Episode Tracker features, so that a history of their information is in a single place:

> The way the Trigger Tracker is set-up I don’t find that helpful at all. I think unless it interfaces with the episode tracker, for me, I don’t see how it’s helpful...There’s no relevance to an episode. [Participant 7]

**Appointment Reminder**

Participants liked the idea of an Appointment Reminder tool where they can keep track of AFib and other medical appointments. However, most participants already used other tools to keep track of their schedule and seemed unlikely to adopt the AFib Connect Appointment Reminder unless there is an easy way to sync appointment information between their personal calendar or electronic medical record.

**News Feed**

Participants tended to rate the News Feed feature lower than the app’s other features. Although some participants liked the idea of having access to the latest heart health-related research and information, many felt the News feed content was not tailored enough to their specific needs.

**App Overall**

Participants’ perceptions of the app’s overall usefulness can be organized into three key themes: the needs of the patient segment, how well the app’s design supports the patient workflow, and whether the app can support the patient over time.

Although the target population of this study was any nonvalvular AFib patient who has been prescribed an NOAC, our data identified a few distinct patient segments within this group: (1) newly diagnosed patients versus those who have been managing AFib for an extended period and (2) patients who are otherwise healthy versus those with multiple comorbidities. As all the participants in our study had been managing AFib for more than a year, they expressed that, overall, the AFib Connect mobile app would have been significantly more useful when they were first diagnosed, still learning about AFib, and still experiencing symptoms:

> Again, it's more for the medication, appointment reminders, maybe trigger tracker kind of things because most of the other stuff I've been through and so I have a pretty good knowledge of the condition now. So I don't think I'd be going back to the library much. I'm not going to obviously go back to the AFib guide, nor the heart rate monitor. [Participant 10]

Similarly, although our study participants were relatively young and had few health conditions apart from AFib, previous research has shown that on average AFib patients are older and have multiple comorbidities [1,19]. We assume that older individuals and those managing multiple conditions would identify different needs for the app than the sample in our study.

Participant feedback also indicated that the app’s overall usefulness is impacted by how well the designed path through the app matches their natural workflow. For example, some participants suggested having the Heart Rate Monitor, Trigger Tracker, and Episode Tracker data and features interface more seamlessly to more easily track key data. One participant also suggested syncing the app’s data to their online medical record so that all their AFib health information can be accessed from a single place:

> I don’t know if there’s ever going to be a way to get it so that it interfaces or interacts with the [online medical record] so that maybe you come in and you log in and then somehow it ties into all your information that's there. [Participant 10]

Study data also highlighted how patient perceptions of the app may shift over time. Participants discussed in detail how their needs now differ greatly from when they were first diagnosed and still experiencing symptoms. Similarly, we expect that as an AFib patient gets older, or experiences an improvement or
deterioration in their condition, that their health priorities and needs will also change [20]. Whether the app can continue supporting a patient’s evolving needs will greatly impact its overall usefulness.

**Discussion**

**Principal Findings**

Results from this study provided greater insight into patient use and acceptability of the AFib Connect mobile app. Additionally, it helped paint a more complete picture of the everyday experience of AFib patients. By giving participants the opportunity to use the app in their natural environment and using qualitative research methods to explore perceptions of usability and usefulness, we obtained valuable feedback on its key features, navigation, content, and workflow that can be used to improve the overall design.

Additionally, study results illustrate three design principles that can be applied more broadly across the development of patient health apps: understand exactly who you are designing for, understand the patient’s natural workflow, and understand how patient needs change over time.

**Understand Exactly Who You Are Designing For**

This study demonstrates how the needs of patients can vary depending on how long they have been managing their condition and whether they have additional comorbidities. Understanding what patient segments exist within the larger population of individuals who share a medical condition and designing for their unique needs is essential to building useful and usable apps. Feedback from all participants indicated that the app would have helped them manage their AFib care and treatment in general and significantly more so when they were first diagnosed.

**Understand Patient Workflow**

Qualitative feedback from participants also revealed exactly how features will be used, when they will be used, and how this app will fit within the broader ecosystem of tools and information resources patients access to manage AFib. It became clear after several interviews that participants might have a more streamlined experience if some app features were combined to better reflect the natural workflow of AFib patients. By considering not only the usefulness of each individual feature, but how these features work together to support self-care and treatment adherence, is key.

**Understand How Needs Change Over Time**

Results from this study also indicate that the usefulness of a health app often changes over time, largely based on a patient’s changing health status and knowledge of their condition. Similarly, we suspect a user’s interaction with a health app may evolve the longer they have used the app and are familiar with its content. For a health app to continue to be useful over a long period, its design will need to consider and adapt to the changing needs of its users.

To build more useful and usable tools for self-care and treatment adherence, it is essential to holistically examine the context in which patients experience their condition. We should evaluate whether the app truly meets the core needs of the target population, if it fits into their natural workflow and with the tools they already use, and whether it can continue to provide support throughout each stage of their condition.

**Limitations**

This exploratory pilot study has a few limitations. A key limitation was the study sample. Participants were younger than the typical AFib population [1], all Caucasian, and none were newly diagnosed. Additionally, as most of the participants did not experience episodes during the study, much of the usefulness feedback we received was based on how the app’s features would have been useful in the past, and thus, responses may be influenced by recall bias. Although the results of exploratory studies are not intended to be generalizable, additional future research utilizing a larger and more diverse sample of newly diagnosed AFib patients will be helpful for understanding the app’s wider applicability.

Another limitation was the duration of the study. Although this study was useful for gaining insight into the AFib patient experience and perceptions of the app over a 4-week period, it would be valuable to see if these perceptions change over a longer period of naturalistic use, as patient health conditions and needs evolve.

**Comparison With Prior Research**

Mobile tools to support self-care and medication adherence have previously shown promise in supporting the patient management of chronic conditions [20]; however, this study is the first we are aware of that specifically examines how these tools might be useful for AFib.

The benefits of qualitative methods for gaining rich insight into the real-world use and acceptability of health apps are well documented [21,22]. Additionally, the value of incorporating patient perspectives during the early stages of design and testing of a new intervention is supported by a growing body of research [23,24]. Yet, relatively few medical studies use qualitative research methods to examine patient perceptions of an intervention at an early stage, or at all, before implementation [23]. This can result in less than optimal, or even negative outcomes for patients who receive the intervention [24].

**Conclusions**

The results of this study suggest that mobile tools that target self-care and treatment adherence may be helpful to AFib patients, particularly those who are newly diagnosed. Additionally, participant feedback provided insight into the varied needs and health experiences of AFib patients, which may improve the design and targeting of the intervention.

Pilot studies that qualitatively examine patient perceptions of usability and usefulness are a valuable and often underutilized method for assessing the real-world acceptability of an intervention [25,26]. Additional research evaluating the app over a longer period and including a larger, more diverse sample of AFib patients will be helpful for understanding whether the AFib Connect mobile app and similar tools can be more widely useful.
By expanding our understanding of the AFib patient experience, we can continue to improve the app’s usability and usefulness and its capability for supporting long-term self-care and treatment adherence.

Acknowledgments
This study was sponsored by Daiichi Sankyo, Inc. Partners Connected Health independently oversaw the study design, participant recruitment, data collection, analysis, and reporting.

Special thanks to Lauren Cortese, Sharon Odamey, and Nina Schussler for supporting participant recruitment; Sara Bersche Golas for assisting with quantitative data analysis; and Timothy M Hale, PhD, and Simone Orlowski for providing feedback on the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
A sample of questions from the Week 2 Semistructured Interview Guide.

References
15. Tran K, Leblanc K, Valentine A, Kavanagh D, Zahr N, Ivers NM. Evaluating the usability and perceived impact of an electronic medical record toolkit for atrial fibrillation management in primary care: a mixed-methods study incorporating...
human factors design. JMIR Human Factors 2016 Feb 17;3(1):e7 [FREE Full text] [doi: 10.2196/humanfactors.4289] [Medline: 27026394]


Abbreviations

AFib: atrial fibrillation

NOAC: nonvitamin K antagonist oral anticoagulants

©Jaclyn Hirschey, Sunetra Bane, Moussa Mansour, Jodi Sperber, Stephen Agboola, Joseph Kvedar, Kamal Jethwani. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 15.03.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
The Perceived Ease of Use and Usefulness of Loop: Evaluation and Content Analysis of a Web-Based Clinical Collaboration System

Allison M Kurahashi, MEd; Jennifer N Stinson, RN-EC, CPNP, PhD; Margaret van Wyk, RN; Stephanie Luca, HBSc, MA; Trevor Jamieson, MBI, MD; Peter Weinstein, BEng, MHSc; Joseph A Cafazzo, PEng, PhD; Bhadra Lokuge, MHS; Eyal Cohen, MSc, MD; Adam Rapoport, MHSc, MD; Amna Husain, MPH, MD

The Temmy Latner Centre for Palliative Care, Sinai Health System, Toronto, ON, Canada
Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada
Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
Division of General Internal Medicine, St. Michael's Hospital, Toronto, ON, Canada
Institute for Health System Solutions and Virtual Care, Women's College Hospital, Toronto, ON, Canada
Department of Medicine, University of Toronto, Toronto, ON, Canada
Healthcare Human Factors, Techna Institute, University Health Network, Toronto, ON, Canada
Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada
Department of Paediatrics, University of Toronto, Toronto, ON, Canada
Emily's House Children's Hospice, Toronto, ON, Canada
Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Amna Husain, MPH, MD
The Temmy Latner Centre for Palliative Care
Sinai Health System
60 Murray Street, 4th Floor, Box 13
Toronto, ON, M5T 3L9
Canada
Phone: 1 416 586 4800 ext 7886
Fax: 1 416 586 4804
Email: amna.husain@sinahealthsystem.ca

Abstract

Background: Patients with complex health care needs require the expertise of many health care providers. Communication, collaboration, and patient-centered care positively impact care quality and patient outcomes. Few technologies exist that facilitate collaboration between providers across settings of care and also engage the patient. We developed a Web-based clinical collaboration system, Loop, to address this gap. The likelihood of a technological system’s uptake is associated with its perceived ease of use and perceived usefulness. We engaged stakeholders in the conceptualization and development of Loop in an effort to maximize its intuitiveness and utility.

Objective: This study aimed to report end users’ perceptions about the ease of use and usefulness of Loop captured during usability tests of Loop.

Methods: Participants represented three user types (patients, caregivers, and health care providers) recruited from three populations (adults with cancer, adolescents and young adults with cancer, and children with medical complexity). We conducted usability testing over three iterative cycles of testing and development in both laboratory-based and off-site environments. We performed a content analysis of usability testing transcripts to summarize and describe participant perceptions about the ease of use and usefulness of Loop.

Results: Participants enjoyed testing Loop and were able to use the core functions—composing, posting, and reading messages—with little difficulty. They had difficulty interpreting certain visual cues and design elements or the purpose of some features. This difficulty negatively impacted perceived ease of use but was primarily limited to auxiliary features. Participants
predicted that Loop could improve the efficiency and effectiveness of communication between care team members; however, this perceived usefulness could be compromised by disruptions to personal workflow such as additional time or task requirements. **Conclusions:** Loop was perceived to have value as a collaboration system; however, usability testing findings indicate that some design and functional elements need to be addressed to improve ease of use. Additionally, participant concerns highlight the need to consider how a system can be implemented so as to minimize impact on workflow and optimize usefulness.

(JMIR Hum Factors 2018;5(1):e2) doi:10.2196/humanfactors.7882

**KEYWORDS**

patient-centered care; patient participation; chronic disease; communication; internet communication tools; usability testing; interdisciplinary communication; health communication; continuity of patient care; patient care team; inventions

**Introduction**

**Background**

Patients with chronic diseases have complex needs that require the expertise of many health care providers (HCPs) from different disciplines, institutions, community organizations, and settings of care [1]. Care plans that are collaboratively developed and transparent are critical to the management of patients with complex care needs across the life span [2,3]. Effective communication in teams is essential to achieve coordinated, continuous care [4,5] and has been associated with enhanced patient safety, better patient outcomes, fewer medical errors, and a reduction in health care redundancies [6-9]. Furthermore, the ubiquitous call for patient-centered care includes engaging patients in medical decisions about their care, which can help to improve their understanding of their health care needs as well as their adherence to care plans [2,10-12].

HCP communication and coordination about goals of care across hospitals and community settings can be a major challenge. Patients have reported dissatisfaction because of poor information exchange between providers, low levels of active patient engagement, and insufficient coordination at time of transition [6,13,14]. At these instances of collaborative breakdown, uncertainty about roles and fragmentation of care [15] results in families taking on responsibilities as communication intermediaries between multiple providers [1,4,16-18]. Furthermore, fragmented care is associated with more frequent emergency department visits, decreased functional status, and higher costs associated with care [19-21].

**Existing Technologies**

Technologies such as mobile texting, email, and messaging systems have been found to positively impact communication between HCPs [22] and may be effective in facilitating coordination of patient care [2,23,24]. Other health information technology (HIT) such as shared electronic health records (EHRs), personal health records, Web-based communities and learning resources, and telehealth also show potential for improving patient care [25,26]. Reviews suggest that improved access to information via HITs can foster patient engagement and empowerment by improving their health information competence, informed decision making, communication with HCPs, and control over their care experiences [4,25-33]. Furthermore, providers may better understand their patients’ needs [34]. Secure patient-provider messaging via patient portals and EMRs also demonstrated successful uptake [35-39] and was perceived by patients and providers to improve communication and information flow [40].

However, these interventions have limitations: giving patients access to their medical information does not guarantee that they will understand that information [34]. In addition, EMRs are often restricted by organizational boundaries, thus inhibiting a longitudinal understanding of a patient’s health [34] and collaboration across sites. Although improved access to information supports individual decision making, without means for interactive discussion, these interventions may fall short of promoting shared decision-making [24]. One study that evaluated system use and user experiences of Web-based communication between patients and their interprofessional care teams reported improved accessibility, efficiency, and transparency [41]. There are few other, if any, studies that investigate existing communication technologies that simultaneously promote collaboration across organizational boundaries and engage the patient in their care [42].

**Development of Loop**

We developed a clinical collaboration system to address these gaps at the intersection of clinical care and information technology. The Web-based system, which we call Loop, provides a secure environment for individual providers to assemble as a team with their patient for care-related communication and collaboration [42]. Each team, or Patient Loop, can include the patient, one or more of their caregivers, and various HCPs involved in their care (Figure 1).

We employed a user-centered design (UCD) [43,44] approach to engage stakeholders (ie, clinicians, researchers, designers, developers, and end users) as early engagement is believed to promote greater uptake at implementation [26,45]. In particular, we used ethnography [46], affinity diagramming [47], cooperative prototyping, and dramatic simulation [48] to involve stakeholders in the conceptualization and generation of system requirements for Loop. Through this process, we defined the following two usability objectives to guide the development of Loop: (1) Loop will be intuitive and easy to use, requiring minimal or no instruction, and (2) Loop will be useful in the care of patients with complex care needs. We continued to engage stakeholders during the development and refinement stages via usability testing, prototyping, and pilot testing. Usability testing involves a representative sample of intended end users interacting with a system to generate insights that will optimize the design and user experience [49-51].
The technology acceptance model (TAM) is commonly used to evaluate new technology systems [52,53]. TAM aims to predict a user’s acceptance and adoption of an information technology system based on two constructs: perceived usefulness (PU) and perceived ease of use (PEOU) [52]. PU evaluates how a tool might enhance job performance, effectiveness, and productivity [52,54,55]. PEOU assesses the perceived effort required to learn and interact with a tool [52,55]. The more useful and easier to use a tool is perceived to be, the more likely it will be accepted [52]. Original applications of TAM employed quantitative metrics; however, qualitative interviews have also confirmed that PEOU and PU are main factors affecting intention to use and explored what is meant by these complex terms [56]. Using TAM as a lens for the qualitative analysis of usability testing experiences, we aimed to understand how end users perceive Loop’s ease of use and usefulness.

Methods

Study Design

The data collected during the iterative cycles of qualitative usability testing were used to evaluate and improve the Loop prototype. Once all data were collected and the prototype completed, a descriptive content analysis of a subset of usability testing data was performed to determine PU and PEOU.

Clinical Collaboration System

The core functionality of Loop includes (1) composing and posting messages that are visible to team members who are part of that Patient Loop (Figure 2) and (2) viewing messages on a central Message Stream (Figure 2). Loop also includes the following auxiliary features: (1) tagging messages with specific labels, which can then be used for filtering messages using the Issues feature (Figure 3); (2) tagging specific team members so that they will receive an email notification about the new message using the Attention To feature (These messages are visible to the entire team on the Message Stream; Figure 4); and (3) selecting whether a message will be visible to the whole Loop (patient, caregiver(s), and HCPs) or only between HCPs on the team using the Team Only feature (Figure 5). The need for these features was identified in the earliest stages of conceptualization. As such, some version of the Issues, Attention To, and Team Only features appeared in all Loop prototypes. Our earlier publication describes usability testing and participant perceptions of the Team Only feature. We found that all participant types endorsed the inclusion of a separate view for HCP-only communication [42].

Sample

A total of 89 participants completed usability testing: 23 patients, 19 caregivers, and 47 HCPs. Two patients and 2 caregivers completed usability testing together. Participants were recruited from the following three populations: adolescents and young adults with cancer (AYAC) [57], adults with cancer, and children with medical complexity (CMC) [3]. These represent populations across the life span with multiple comorbidities that have high service needs and require care from multiple providers across health care settings. AYAC participants were patients aged between 15 and 25 years and receiving oncological care (n=15). Adult cancer participants were patients (n=8) and caregivers of patients (n=12) receiving oncological care. CMC participants were caregivers of children with severe functional limitations (n=7). Convenience samples of eligible participants were identified by HCPs at three academic institutions specializing in cancer, palliative, and pediatric care located in Toronto, Canada (Princess Margaret Cancer Centre, The Temmy Latner Centre for Palliative Care, and The Hospital for Sick Children, respectively).
Convenience sampling is an accepted form of sampling in usability testing [50]. HCPs from various disciplines who work with oncology and CMC patients were purposefully recruited from these institutions and other community-based care organizations to maximize representation across population, provider type, and clinical environment (AYAC: n=16; adult cancer: n=19; and CMC: n=12). The study protocol was approved by the research ethics boards at each site, and informed consent was obtained from all participants.

Usability Testing Protocol

Usability testing was completed over three rounds of iterative laboratory-based testing, which included a desktop computer in a quiet room equipped with microphones, video cameras, and one-way mirrors. This was supplemented with concurrent off-site testing, which was conducted with a separate sample of participants at a convenient location, such as their office, home, or meeting rooms. Each round of testing evaluated a prototype of increasing fidelity: low (iPad and paper prototype), medium (wireframe), or high (working prototype). The high-fidelity prototype was tested and refined over a series of rapid development sprints. Participants completed questionnaires collecting demographic characteristics and technology use before testing.

Participants were presented with a set of standardized task-oriented scenarios targeting new or refined features, processes, or design elements. Table 1 lists the categories of tasks completed by the different participant types. Not all features were tested with every participant type; however, participants may have interacted with features during exploration or while completing another task. Scenarios were worded to allow for a natural flow of interactions rather than as step-by-step instructions (eg, “Your pain is now being well controlled by your medications. How would you alert the team?”). Using this approach, facilitators observed the functions participants used and how they used them as well as common navigation errors and inefficiencies. Participants were asked to think aloud and verbalize their choice of actions while interacting with the system. To supplement the think aloud component, facilitators asked questions throughout the testing session to capture users’ reactions and reflections (eg, “Do you have any general thoughts about the layout of the message area? How did you find replying to a message?”). All usability testing sessions were audio-recorded and transcribed. Screen captures of testing sessions were recorded but were not used for this analysis.

Data Analysis

Demographic data were analyzed using Microsoft Excel (Microsoft Corporation, Redmond, Washington) to determine categorical frequencies. Medians were calculated for participant ratings of comfort using various technologies.

Content analysis was used to analyze the usability testing interview transcripts and generate a descriptive summary of
users’ PEOU and PU of Loop [58]. Three reviewers (BL, AK, and MVW) independently open coded transcripts in NVivo version 10 (QSR International, Burlington, MA). Codes were hierarchically organized and served as the coding framework, which was continually adapted and reviewed with two senior team members (AH and JS) at key points in the analysis process. This review and coding process familiarized reviewers with the data and informed focused in-depth analysis. Open codes were then categorized as PEOU or PU. Themes within these categories were identified using questions derived from the key constructs of TAM [59]:

PEOU: “What elements were easy or difficult to use?” and “Why were they perceived to be easy or difficult?” Participant responses were analyzed to identify comments that explicitly expressed positive or negative sentiments about the ease of navigating Loop, or sentiments of confusion, frustration, or satisfaction from which ease or difficulty could be inferred.

PU: “What would people use Loop for?”, “Why would that be useful?”, and “How could that improve care?” During the analysis of these responses, additional questions of “What are the factors that impact Loop’s usefulness?” and “What strategies could mitigate the barriers?” were also explored.

Usability testing yielded 87 transcripts. We used a maximum-variation sampling [60] approach to maximize the representation of perspectives and experiences across the following categories: population type (adult cancer, AYAC, and CMC), laboratory-based or off-site testing, user type (caregiver, patient, and HCP), type of HCP, and version of Loop tested. On the basis of the analysis of this subsample of transcripts, all themes became saturated and no new concepts emerged from the data, prompting the decision not to code any further transcripts. A total of 48% (42/87) of transcripts were included in analysis.

**Figure 3.** High-fidelity prototype demonstrating the Tag Issues feature being used to apply tags while composing a message and in the message stream (top), and to edit or update issue status in the filtered view (bottom).
Figure 4. High-fidelity prototype version of the Attention To feature being used to tag team members when composing a message and resulting visual cues.

When composing a message, users can select team members who should be notified about the posted message.

Team members who have been notified appear in the Attention To list.

Messages marked Attention To “me” have yellow visual cues.

Figure 5. High-fidelity prototype version of the Team Only feature being used to set visibility while composing a message and resulting visual cue.

HCP: health care provider.

When composing a message, HCP users can use the Team Only feature to set who will be able to see the message in the message stream.

Messages with a blue ring and icon indicate that it is not visible to patients or caregivers, and only visible to other HCPs.
Table 1. Participant tasks and participant types who completed each task.

<table>
<thead>
<tr>
<th>Participant tasks</th>
<th>Patient</th>
<th>Caregiver</th>
<th>Health care providers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Navigating outside a Patient Loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Register for Loop and set up a profile</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Log in to Loop</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Find a patient in the patient list</td>
<td>--</td>
<td>--</td>
<td>X</td>
</tr>
<tr>
<td>Create a new Patient Loop</td>
<td>--</td>
<td>--</td>
<td>X</td>
</tr>
<tr>
<td><strong>Navigating within a Patient Loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explore a Loop</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Messages</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read messages</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>View conversation</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Filter messages by issue</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Compose a new message</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Use the attention to feature</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Tag an issue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create a new issue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Send a team only message</td>
<td>--</td>
<td>--</td>
<td>X</td>
</tr>
<tr>
<td>Reply to a message</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Update an issue status and summary</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Manage the team</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Find team member information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invite a new team member to a Patient Loop</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Results**

**Participant Characteristics**

All adult cancer patients were older than 50 years, and all AYAC patients were younger than 30 years, representing the oldest and youngest participants tested (Table 2). Caregivers and HCPs were concentrated around the central age ranges; 95% (18/19) of all caregivers and 96% (44/46) of all HCPs were aged between 30 and 69 years. A majority of participants had access to computers and Internet at home (Table 3). Adult cancer patients were the least comfortable of all participants using the surveyed technologies, and AYAC patients were the most comfortable (Figure 6). Of all technologies, participants in all populations except AYAC were least comfortable using social media.

**Qualitative Findings**

During usability testing sessions, participants provided feedback about the PEOU and PU of the system. Themes were identified within each of these categories (Figure 7). Broadly, analysis revealed two types of problems that negatively affected the PEOU of Loop: visual design problems and incorrect use of features. With regard to PU, participants felt that Loop could improve the efficiency and effectiveness of communication in real-world use. However, PU could be negatively affected if Loop disrupted individual workflow. Participants suggested features to mitigate Loop’s potentially disruptive impact and improve ease of use. These findings are described in detail below. Responses were consistent across user type (patient, caregiver, and HCP) and populations (CMC, AYAC, and adult cancer) unless otherwise noted.

**Perceived Ease of Use**

The majority of participants enjoyed testing Loop. They felt that the layout and design were easy to navigate and that the core functions (composing, posting, and viewing messages) were intuitive to use. One participant stated the following:

> Very clear. I like the layout, it’s very simple. It doesn’t have a lot of like sub-links and things flashing that distracts one’s attention. [Caregiver, CMC, ID#18, high-fidelity prototype]

Another participant stated the following:

> I do like that it’s very clean and there isn’t a lot of information, it isn’t very busy. So yeah, overall I really like it, the format. [HCP, CMC, ID#36, high-fidelity prototype]

**Factors Negatively Impacting Loop’s Perceived Ease of Use**

Negative feedback about ease of use was mostly related to the visual design and use of auxiliary features such as Attention To, Issues, and specific visual cues in the Message Stream. We observed that participants were less likely to volunteer
comments if they did not encounter a problem when navigating the system.

**Visual Design Problems**

Some participants did not perceive icons as clickable, did not notice visual cues, or had difficulty interpreting the meaning of icons or visual cues. These errors may highlight problems with certain visual design elements. For example, some participants did not identify the clickable icons that would allow them to complete tasks related to replying to messages, viewing conversations, and editing Issue statuses. In other situations, participants were unable to identify what the visual cues were trying to convey or did not perceive these cues at all. In particular, the blue ring and icon on profile pictures indicating messages visible only to HCPs and the yellow background identifying messages directed at the user were often overlooked or misinterpreted. Problems interpreting visual cues did not impair participants’ ability to read messages in the Message Stream, but useful contextual information related to each message may not have been understood. One participant stated the following:

*I’m interested in this pencil down here. Does that mean something?...I have no idea. Maybe like an edit? Maybe that means edit the page?* [Patient, AYAC, ID#01, high-fidelity prototype]

Another participant stated the following:

*So I don’t know what this means though, this circle and then the little person. Is that like consult, is that what that means? I don’t know what that means.* [HCP, CMC, ID#36, high-fidelity prototype]

<table>
<thead>
<tr>
<th>Table 2. Participant demographic information.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Age in years</td>
</tr>
<tr>
<td>10-29</td>
</tr>
<tr>
<td>30-49</td>
</tr>
<tr>
<td>50-69</td>
</tr>
<tr>
<td>70-89</td>
</tr>
<tr>
<td>Education</td>
</tr>
<tr>
<td>High school - current</td>
</tr>
<tr>
<td>High school - completed</td>
</tr>
<tr>
<td>College or university</td>
</tr>
<tr>
<td>Professional or graduate</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Diagnosis</td>
</tr>
<tr>
<td>Lung cancer</td>
</tr>
<tr>
<td>Ovarian cancer</td>
</tr>
<tr>
<td>ALL\textsuperscript{c}</td>
</tr>
<tr>
<td>AML\textsuperscript{d}</td>
</tr>
<tr>
<td>Ewing sarcoma</td>
</tr>
<tr>
<td>Rhabdomyoscarcoma</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
</tr>
<tr>
<td>Osteosarcoma</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

\textsuperscript{a}AYAC: adolescents and young adults with cancer.  
\textsuperscript{b}CMC: children with medical complexity.  
\textsuperscript{c}ALL: acute lymphocytic leukemia.  
\textsuperscript{d}AML: acute myeloid leukemia.
Table 3. Participants’ use of technology.

<table>
<thead>
<tr>
<th>Technology characteristics</th>
<th>Patient Adult cancer (N=8), n (%)</th>
<th>Patient AYAC (N=15), n (%)</th>
<th>Caregiver Adult cancer (N=12), n (%)</th>
<th>Caregiver CMC (N=7), n (%)</th>
<th>Health care providersa Adult cancer (N=19), n (%)</th>
<th>Health care providersa CMC (N=11), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has computer at work or school</td>
<td>5 (63)</td>
<td>13 (87)</td>
<td>10 (91)</td>
<td>6 (86)</td>
<td>19 (100)</td>
<td>11 (100)</td>
</tr>
<tr>
<td>Has computer at home</td>
<td>7 (88)</td>
<td>14 (93)</td>
<td>11 (92)</td>
<td>7 (100)</td>
<td>19 (100)</td>
<td>11 (100)</td>
</tr>
<tr>
<td>Has Internet at home</td>
<td>7 (88)</td>
<td>14 (93)</td>
<td>11 (92)</td>
<td>6 (100)</td>
<td>19 (100)</td>
<td>11 (100)</td>
</tr>
</tbody>
</table>

**Hours on computer per day**

<table>
<thead>
<tr>
<th></th>
<th>Patient</th>
<th>Caregiver</th>
<th>Health care providersa</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>2 (25)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>1-7</td>
<td>5 (63)</td>
<td>12 (80)</td>
<td>8 (67)</td>
</tr>
<tr>
<td>&gt;7</td>
<td>1 (13)</td>
<td>3 (20)</td>
<td>3 (25)</td>
</tr>
</tbody>
</table>

**Hours on Internet per day**

<table>
<thead>
<tr>
<th></th>
<th>Patient</th>
<th>Caregiver</th>
<th>Health care providersa</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>2 (25)</td>
<td>1 (7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>1-7</td>
<td>6 (75)</td>
<td>12 (80)</td>
<td>10 (83)</td>
</tr>
<tr>
<td>&gt;7</td>
<td>0 (0)</td>
<td>2 (13)</td>
<td>1 (8)</td>
</tr>
</tbody>
</table>

These data were not collected for AYAC health care providers.

bAYAC: adolescents and young adults with cancer.

cCMC: children with medical complexity.

Figure 6. Participant comfort using various technologies. Rating scales from 0 (do not use) to 4 (very comfortable). Adult cancer patients had a median rating of 0 for mobile phone comfort. AYAC: adolescents and young adults with cancer; CG: caregiver; CMC: children with medical complexity; HCP: health care provider; PT: patient.
Participants suggested adding cursor labels to describe the purpose or indicate clickability of icons and visual cues. Cursor labels (text that appears when a user hovers the cursor over an icon) were introduced in the high-fidelity prototype of Loop and observed to improve navigability when tested.

For most visual design problems, participants indicated that subsequent use of Loop would be easier after they received orientation and instruction about Loop’s icons and visual cues.

Incorrect Use of Features

Some participants incorrectly used the Attention To and Team Only features when composing messages. For example, some participants incorrectly used the Attention To feature to select team members to whom the message would be visible. HCPs sometimes perceived a redundancy between the Attention To list and Team Only toggle (a feature only available to HCPs), despite these features controlling different things, as shown in a conversation below:

Interviewer (I): Do you expect that the other team members would be able to see that message as well?
Respondent (R): I wouldn’t expect it if I didn’t select any [in the Attention To feature].
I: Okay. So you expect only Dr. Torres would be able to see the message?
R: Yes, unless I sent it to all of her team, then I would click everybody that I want to see the message.
[Caregiver, CMC, ID#18, high-fidelity prototype]

Another participant stated the following:

Given that [the Team Only toggle is set to] patient and team, the patient shows up again in this list [Attention To list], which seems redundant because if it’s going to the patient anyway, then why have it twice. [HCP, AYAC, ID#01, high-fidelity prototype]

This feedback will be addressed in future development of Loop with additional rounds of usability testing.

Perceived Usefulness

Participants indicated that Loop’s PU lies in its potential to improve efficiency and effectiveness of communication about patient care within a team. Participants also highlighted that PU could be negatively affected if individual workflow is disrupted.

Positive Impacts on Communication

Improving Efficiency

Participants across stakeholder groups felt that a central communication space such as Loop could improve the efficiency of communication about patient care within a team. The ability to post a single message that is viewable by all team members could save patients and caregivers the time and frustration of repeating information to multiple members of their care team. This ability to access multiple providers with a single post would also be useful when patients or caregivers are not sure to which HCP they should direct their questions or updates. One participant stated the following:

One of the things that I see as being useful about this is it should cut down the amount of time that family members, patients and caregivers are spending repeating information. [Caregiver, adult cancer, ID#09, low-fidelity prototype]

Another participant stated the following:

And I think as far as the patient, instead of them sending you an email, if they send something on that system, at least everybody can kind of contribute to it, if there a concern she’s having. [HCP, AYAC, PMH, ID#01, low-fidelity prototype]

Participants indicated that they currently receive information from a variety of technologies, such as email, paging, and text...
messaging, and Loop could be useful for consolidating incoming messages in one place. This was primarily voiced by HCPs but also mentioned by some patients. One participant stated the following:

...rather than posting and copying and pasting to multiple doctors, emailing them...and getting their opinion on it—I think it provides an easier and quicker way to get in contact with everybody. [Patient, AYAC, ID#03, high-fidelity prototype]

Another participant stated the following:

This would be a forum where we’re all connected, whereas email, there’s one here, one there, people are doing different things with the patient but not necessarily communicating in one forum. [HCP, CMC, ID#37, high-fidelity prototype]

Patients and caregivers also suggested that being able to post a symptom update or question as it occurs, even if they are not expecting an immediate response, could be more efficient than remembering to ask the question or recall a symptom at their next appointment. One participant stated the following:

I don’t need to wait to call them. And as soon as I have any questions, I can open up that thing and write it down, my questions. And I can get the answers as soon as possible. So, it’s a really good communication thing. [Caregiver, CMC, ID#32, high-fidelity prototype]

**Improving Effectiveness**

Participant responses indicated that a central communication space could also improve the effectiveness of communication; all team members could be aware of what is going on with patient care even if they are not directly involved at that time. One participant stated the following:

I think that the whole purpose [of this system] is to have everybody within the team to know all the information about me, the patient. [Caregiver, adult cancer, ID#21, high-fidelity prototype]

Another participant stated the following:

I think it really does facilitate people knowing what’s going on with patients who have multiple providers. [HCP, CMC, ID#24, high-fidelity prototype]

Patients and HCPs alike felt that this increased transparency would promote patient-centeredness, patient engagement, and coordination between providers.

**Patient-Centeredness**

Patients and caregivers felt that including the whole team in discussions about care could broaden HCPs’ understanding of the patient’s needs beyond a specific specialty. HCP participants did not explicitly comment on this. One participant stated the following:

I know my mom has her palliative doctor, her radiation oncologist and a urologist that are all helping in her care. So, just to have each person have a full understanding...not just their specialty, but a broader understanding, it could be great. [Caregiver, adult cancer, ID#08, low-fidelity prototype]

All participants suggested that Loop could provide patients with a way to contribute to the conversation, take an active role in decision making, and understand how decisions are made. One participant stated the following:

I’m interested in like seeing like, you can actually see the doctor’s thought process and so many times when you’re in an office, you don’t get to see that. [Patient, AYAC, ID#01, high-fidelity prototype]

Another participant stated the following:

I think this is amazing from a transparency standpoint. That you know, that the patient’s seeing all of the discussions going on and who’s involved in their care. [HCP, AYAC, ID#06, high-fidelity prototype]

**Coordinated Care**

Caregivers of adult and CMC patients indicated that they are typically responsible for synthesizing, reconciling, and relaying information between individual providers. They believed that Loop could improve collaboration and cooperation directly between providers, thus decreasing the burden of coordination on the caregivers. Although HCPs did not comment specifically on caregiver burden, they did acknowledge that Loop could facilitate communication between providers, patients, and families across locations to promote collaboration on things such as discharge plans. One participant stated the following:

I really think it’s a great idea because there’s no question that certainly in our experience when you start having multiple doctors involved with multiple areas of specialty, it’s a challenge to keep things coordinated for sure. I think this is going to be a very, very helpful tool for everybody concerned, both to the team and the patients. [Caregiver, adult cancer, ID#13, medium-fidelity prototype]

Another participant stated the following:

We always struggle with communication between teams. And sort of, in terms of everybody being on the same page. So I think something like this would be great. [HCP, AYAC, ID#07, high-fidelity prototype]

**Factors Negatively Affecting Perceived Usefulness**

Participants described several factors that may add tasks or time to their workflow and, consequently, would negatively affect the PU. Participants also identified or alluded to several features that would help to mitigate the negative impact of the factors outlined below.

Across all stages of prototyping, patient, caregiver, and HCP participants felt that including patients and caregivers on all messages would reduce the efficiency of communication within Loop.

In response to this consistent feedback, the option to restrict message visibility to user-specified subteams was introduced.
in the low-fidelity prototype and further refined as the Team Only setting in the medium-fidelity prototype.

Despite the introduction of the Team Only feature, participants continued to express a tension between a need for efficiency and wanting to maintain patient engagement and transparency. One participant stated the following:

But if I’m including [the patient and caregiver] in the message, I have to think about the language more than if I were just including the team. So, that’s raising the issue for me of, is it easier for me to not include [the patient and caregiver] in every [message]? I would rather have [the patient] as part of the team. But I can see that it’s a little more of a challenge. [HCP, adult cancer, #14, medium-fidelity prototype]

HCP participants also expressed concern that failing to integrate Loop with other systems such as EHRs could reduce efficiency if they are required to document or search for information in multiple systems. One participant stated the following:

The only thing I worry about is information in two different places. It’s the information in the chart and information here... and just both of those pieces of information are a big process. [HCP, AYAC, ID#06, high-fidelity prototype]

As a pragmatic workaround to a multiple EHR environment, we introduced a feature in the high-fidelity prototype that exports messages as a PDF for upload into an EHR.

HCP participants felt that high volumes of messages could reduce efficiency by making it difficult to find information that is relevant and has high priority. One participant stated the following:

I do think there’s a possibility of having many, many messages that are totally irrelevant to certain members of the team and then having the whole page be things that aren’t necessarily [relevant]. [HCP, adult cancer, ID#17, high-fidelity prototype]

Filtering was described as a way to make specific or relevant messages easier to find by reducing the number of messages one has to sort through. The ability to filter messages in Loop was first introduced in the low-fidelity prototype and was fully functional in the high-fidelity prototype. Messages can be filtered in a number of ways: by threaded conversations, by issue (Issues feature), by messages directed at me (Attention To feature), by messages flagged by the user (Starred Messages feature), or by sender. One participant stated the following:

Filtering messages is good. I think that’s important because, this, over time is going to be enormous. [Caregiver, CMC, ID#20, high-fidelity prototype]

Another participant stated the following:

I think it’s great. I think the fact that you can...zero in on the particular issues is really important, because I suspect some of these can go on for weeks and months. [HCP, AYAC, ID#06, high-fidelity prototype]

All user groups were concerned about posting messages in Loop that are unread by the intended person. As a result of these missed messages, decision makers may have incomplete or fragmented information. One participant stated the following:

Yeah, the worry I have [is that]...you do this for six months and you realize one of them just never looks. And then, you’re like, now what? How do I ring that person’s bell? Do I have to go back to conventional means and use the phone? [Caregiver, CMC, ID#20, high-fidelity prototype]

Another participant stated the following:

...you don’t know how often someone is going to be checking this. They’re probably checking [Loops] on an as needed basis and so there is potential to be missing messages. [HCP, CMC, ID#34, high-fidelity prototype]

Participants suggested that notifications alerting team members about relevant messages and prompting them to log in would be a useful feature. This was felt to be especially important after long periods of inactivity when messages are more likely to be missed. The Attention To feature, which generates email notifications, existed in all prototype fidelities; however, as described in the PEOU section, some participants were not sure how to correctly use this feature. One participant stated the following:

I would [want notifications], yes, just because I think it’s helpful to have it flagged rather than to have to just go back and continually check. [Patient, adult cancer, ID#13, medium-fidelity prototype]

Another participant stated the following:

But for some of them you don’t hear from them for months so it may be helpful for a notification that there’s a new message or something on whichever kid it’s on. [HCP, CMC, ID#24, high-fidelity prototype]

Discussion

Principal Findings

This study evaluated end users’ PEOU and PU of a Web-based clinical collaboration system, Loop. During usability testing sessions, patients, caregivers, and HCPs were able to accomplish tasks testing the core functions of Loop, including viewing, composing, and posting messages. Participants had difficulty interpreting certain visual design elements and using auxiliary features. In these instances, participants were unable to navigate certain features as intended; however, most participants were able to understand features after a brief period of exploring Loop. With regard to usefulness, participants expressed that Loop could be a valuable system for communication between patients, caregivers, and HCPs. Understanding how potential end users perceive the ease of use and usefulness of a technology is important because these factors have been associated with users’ intention to adopt a technology [52,54,61,62]. Furthermore, understanding the underlying causes of difficulty...
associated with features helps in identifying strategies to improve the usability of Loop.

Difficulty using the Attention To and Team Only features highlights the value of adhering to visual design conventions to improve user experience [63-65]. Although the concepts of message notifications and visibility are used in other social media platforms, it is possible that participants’ existing mental models, beliefs about how a system will work based on previous experiences, may have conflicted with the terminology we applied [66].

Older adults reported reduced comfort using most technologies including social media. This is consistent with Pew Research Centre survey data on social media adoption and usage trends from 2013 [67]. Lack of familiarity, complex interfaces, and privacy concerns have been cited as barriers to technology uptake in older adults [68,69]. To make new technologies accessible to older and new users, interfaces should be simple and consistent, and language should be easy to understand and free of jargon that assumes users’ prior knowledge [68,69].

Following a brief explanation by the usability testing facilitator about system features and the introduction of cursor labels, participants felt that Loop would be easier to navigate in subsequent uses. It is possible that embedded instructional features, such as expanding the application of cursor labels to visual cues, may also aid novice navigation and reduce the time, errors, and difficulty associated with task completion [70]. Any instructional elements will need to be unobtrusive for those more familiar with social media conventions.

Participants in this study highlighted Loop’s potential to improve communication and collaboration. We are not aware of any randomized control trials that demonstrate the impact of communication technologies on cross-institutional and interprofessional collaboration or on patient outcomes, such as symptom management, quality of life, length of stay, or mortality rates. Results from a pilot randomized control trial suggest a trend of improved continuity of care with access to Loop [71]; however, this finding must still be confirmed in a full-scale effectiveness trial. Without this type of robust evaluation, it is difficult to predict how the perceived benefits of Loop will translate to real-life use.

Unlike the factors impacting PEOU, which relate to how the system is designed (looks or operates), the factors impacting the PU generally relate to Loop’s perceived impact on participant workflow. Although we do not have data on actual impact, participants in this study predicted a number of workflow disruptions that have been previously noted in other evaluations of eHealth integration into clinical contexts: user frustration about not knowing whether message was received [22], greater quantity of messages [40] resulting in decreased quality of messages [22], duplication of workflow [72,73], altered communication patterns [72,74], and difficulty identifying important information because of abundance of information [74,75]. Several reports have found that actual message volume between patients and HCPs within electronic communication tools tends to be modest [39,76,77], suggesting that provider concerns may be unwarranted. Access to Web-based messaging does not appear to impact the frequency of face-to-face visits [37,38] or telephone and email volume [39,76] but rather supports these typical interactions [26].

Some of the barriers to adoption identified by the participants in this study are not represented by TAM. Indeed, TAM has been critiqued for not considering the impact of external factors, such as user workflow, organizational characteristics, and social context, on users’ acceptance of technologies [54,78,79]. Financial compensation structures, lack of empirical evidence about a system’s usefulness, personal characteristics such as computer experience [78,79], professional conflict [25], and power dynamics [80] are other examples of external factors that act as barriers to system adoption. An updated version of TAM, the unified theory of acceptance and use of technology, includes the variable Facilitating Conditions, which acknowledges the influence of perceived organization and technology infrastructure on uptake [55]. To be successful, any real-world implementation will need to consider a full range of internal and external factors.

Several features to mitigate negative impact on workflow were integrated in later stages of Loop’s prototyping. In some cases, these features were difficult to use or were overlooked entirely. Not surprisingly, system features are only useful if they are easy to identify and use [52,61,75]. Indeed, studies have shown a statistically significant relationship between PEOU and PU of a tool [59], reinforcing the need for ongoing UCD and evaluation throughout the life cycles of development and implementation.

Limitations

Findings of this study are based on participants’ subjective feedback about the system and not evaluation of objective measures, such as actual time to complete tasks or number of errors. Additionally, participant feedback about how they might use the system was based on an interaction with the system guided by a clinical scenario and not in real life. Introducing information technology into complex adaptive health care environments has additional design, development, implementation, and evaluation challenges across a range of domains: hardware and software, clinical data definitions, human computer interfaces, incentives and behavior, workflow and communication, internal policies, external regulations, and the need for ongoing monitoring [81]. Although scenario-based usability testing can address and anticipate many issues in real clinical environments, real-life complexities create a need for ongoing assessments of how tools work across settings and users. Ongoing evaluations of Loop, including a pragmatic randomized trial, continue to assess PEOU and PU as key metrics in parallel to health-related outcomes and systemic factors that impact usage and behavior.

The majority of HCPs who participated in this study were recruited from academic institutions in an urban setting. Recruitment began with the networks of the study investigators, all of whom work in academic institutions, resulting in less representation from community or rural settings. However, we sampled patients, caregivers, and HCPs across different complex care populations, and the feedback was consistent. Usability testing was structured to test specific functions and features that had been introduced or updated during a
developmental iteration. This influenced the functions or features that participants talked about in each cycle of testing. Participants were given tasks but not instructions on how to complete them. It is possible that some of the errors observed may have resulted from misunderstanding the task rather than with the system itself. Using the think aloud approach, participants were more likely to verbalize negative feedback and less likely to comment on easily navigated features when interacting with the system.

**Conclusions**

Loop was perceived to have the potential to improve the efficiency and effectiveness of communication about patient care. Results from usability testing point to the importance of having intended users interact with the system at early stages of development to ensure the system is both usable and useful, thereby increasing chances of system adoption in a real-life setting. A number of issues with the system were anticipatory, concerning potential challenges with integrating the system into real-life environments and workflows rather than proximal usability problems. It is, therefore, essential to continue assessing and enhancing user experience throughout the next phase of research including real-world implementation. Future research should examine the broader sociotechnical characteristics that will influence the implementation and overall benefit of Loop in clinical care.

**Acknowledgments**

The authors would like to acknowledge the following individuals for their contributions to Loop’s development and evaluation: Christopher Obwanga, Fatima Kanji, Renee Desjardins, Russell Goldman, and Cynthia Nguyen. This work was supported by the Ramona Rull Karson Innovation Grant of the Canadian Cancer Society (grant #701156), the Canadian Institutes of Health Research Pfizer Operating Grant in Disease Prevention and Management Competition (FRN 122030), and the Canadian Institutes of Health Research eHealth Innovation Catalyst Grant (CEH 126529).

**Conflicts of Interest**

None declared.

**References**


Abbreviations

AYAC: adolescents and young adults with cancer
CMC: children with medical complexities
EHR: electronic health record
HCP: health care provider
HIT: health information technology
PEOU: perceived ease of use
PU: perceived usefulness
TAM: technology assessment model
UCD: user centered design
Co-Designing a Collaborative Chronic Care Network (C3N) for Inflammatory Bowel Disease: Development of Methods

Michael Seid¹,², PhD; George Dellal², MHA; Laura E Peterson, BSN, SM; Lloyd Provost¹, MS; Peter A Gloor⁴, PhD; David Livingstone Fore⁵, BA; Peter A Margolis², MD, PhD

¹Pulmonary Medicine, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
²JAMES M Anderson Center for Health Systems Excellence, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States
³Associates in Process Improvement, Austin, TX, United States
⁴Center for Collective Intelligence, Massachusetts Institute of Technology, Cambridge, MA, United States
⁵Catabolic Design, Oakland, CA, United States

Corresponding Author:
Michael Seid, PhD
Pulmonary Medicine
Department of Pediatrics
Cincinnati Children's Hospital Medical Center
3333 Burnet Ave
Cincinnati, OH, 45229
United States
Phone: 1 5138030083
Email: michael.seid@cchmc.org

Abstract

Background: Our health care system fails to deliver necessary results, and incremental system improvements will not deliver needed change. Learning health systems (LHSs) are seen as a means to accelerate outcomes, improve care delivery, and further clinical research; yet, few such systems exist. We describe the process of codesigning, with all relevant stakeholders, an approach for creating a collaborative chronic care network (C3N), a peer-produced networked LHS.

Objective: The objective of this study was to report the methods used, with a diverse group of stakeholders, to translate the idea of a C3N to a set of actionable next steps.

Methods: The setting was ImproveCareNow, an improvement network for pediatric inflammatory bowel disease. In collaboration with patients and families, clinicians, researchers, social scientists, technologists, and designers, C3N leaders used a modified idealized design process to develop a design for a C3N.

Results: Over 100 people participated in the design process that resulted in (1) an overall concept design for the ImproveCareNow C3N, (2) a logic model for bringing about this system, and (3) 13 potential innovations likely to increase awareness and agency, make it easier to collect and share information, and to enhance collaboration that could be tested collectively to bring about the C3N.

Conclusions: We demonstrate methods that resulted in a design that has the potential to transform the chronic care system into an LHS.

(JMIR Hum Factors 2018;5(1):e8) doi:10.2196/humanfactors.8083

KEYWORDS
chronic disease; pediatrics; health care delivery; quality improvement
Introduction

Background and Rationale

It has long been known that within the health care system, patients across care settings are prescribed half of indicated care [1-3] and follow through on half of what is prescribed [4]; translating interventions into practice takes too long [5], and research is too expensive, too slow, and does not reflect the needs of patients seen in real-world settings. This is not because of a lack of will or ideas but rather, to the absence of a system in which the efforts and ideas of all stakeholders are translated to improvement.

What if we could create a vastly better chronic illness care system by harnessing the inherent motivations and collective intelligence of patients and families, clinicians, and researchers, so that all could collaborate, at scale, to improve health?

This provocation was the seed for the collaborative chronic care network (C3N) project [6]. The C3N model reflects the Institute of Medicine’s learning health system (LHS) [7] in which health care, improvement, and research are purposefully integrated, but extends that model, via network-based or peer production [8], to all stakeholders. The C3N model reflects scientific advances over the last 20 years in cooperative behavior [9], collective intelligence [10], and organizational architecture for innovation [11], which point to a fundamental principle: people are, by and large, cooperative and generous. The design of systems, including chronic care systems, can hinder or facilitate expression of these impulses.

Design processes are widely used to create and modify products, services, and systems [12]. The purpose of design is to imagine new and better ways to match products, services, or systems with user contexts and goals [13]. Cooperative design, often shortened to codesign [14], refers to actively involving all stakeholders in every stage of the design process, which ensures that the end product meets the needs of all stakeholders. In this way, the design process and the design results become a reinforcing loop, aligning stakeholders and facilitating collaborative action to achieve that design.

Objective

The objective of this study was to report the methods used, with a diverse group of stakeholders, to translate the idea of a C3N to a set of actionable next steps. Although we have previously described elements of the C3N model [6,15], the use of goal-directed design [16], and the formation of a collaborative open-innovation network [17], the unique contribution of this report is describing the process of codesigning—with representatives from all relevant stakeholders and using the idealized design process—an approach that has the potential to transform the chronic care system. We use the case of ImproveCareNow, a learning network to improve health, care, and costs for pediatric patients with inflammatory bowel disease (IBD), as the model case for these methods.

Methods

Human Subjects Protection

This research was reviewed and approved by the institutional review board at the corresponding author’s home institution.

Setting and Population

ImproveCareNow was launched in 2007 to improve health, care, and costs for children with Crohn’s disease and ulcerative colitis; together, IBD [6]. Known previously as PIBDNet, ImproveCareNow originally included nine care centers that used a modified Breakthrough Series model [18] to create a quality improvement (QI) network focused on improving remission rates for their patients [19]. ImproveCareNow initially involved clinicians, without significant involvement of patients, families, or other stakeholders. At the time of the design process (January 2010-July 2011), the ImproveCareNow network had grown to 24 care centers, with data on 2500 patients from 7500+ visits.

The C3N project, funded by an National Institutes of Health (NIH) Transformative Research Awards, aimed to design, prototype, and pilot a C3N—a potentially transformative system for chronic care. The C3N project partnered with ImproveCareNow to help it transform itself from a QI network into a C3N.

The C3N team that led this work was composed of a pediatrician and epidemiologist, a behavioral and social scientist, improvement experts, designers, an expert in collective intelligence, and project management staff. Subject matter experts were integrated into the leadership team: pediatric IBD patients, parents of patients with IBD, pediatric gastroenterologists, and other pediatric IBD clinicians. All members of the leadership team also participated in the design process.

Design Process Participants

The initial design meeting included youth with IBD, parents of children and youth with IBD, pediatric gastroenterologists, nurses and other clinicians, and a variety of other experts including designers, technologists, artists, QI specialists, social scientists, intellectual property experts, and community organizers.

Participants in this initial meeting were recruited from ImproveCareNow by identification of individuals through literature and Internet search, as well as by snowball sampling, in which existing participants are asked to nominate other potential participants.

Design Process

We combined several approaches to design the C3N. These are illustrated in Figure 1. We used theories from the leadership of social movements and collective intelligence to motivate and build cross-stakeholder collaboration, and the idealized design process (specifically, phase 0), including observation, synthesis, and screening, to produce outcomes measures, the design concept, a key driver diagram (KDD; see Figure 2), and potential changes. These are further detailed below.

http://humanfactors.jmir.org/2018/1/e8/
Motivating and Building Cross-Stakeholder Collaboration

We used theories from leadership of social movements [20] to motivate and build cross-stakeholder collaboration. To build motivation, we created forums for patients, parents, and clinicians to share what Ganz calls a public narrative—stories that weave together values and emotion to cause action. We began each design meeting, for example, by having patients and parents share their public narrative. We also developed motivation for cross-stakeholder collaboration by creating a common vision of an idealized state (Idealized Design section, below).

We built cross-stakeholder collaboration by emphasizing the enormity of the challenge and the need for everyone’s expertise and effort using specific messaging around urgency, hope, and self-efficacy to encourage participation (eg, “You can make a difference”) and solidarity (eg, “be part of the solution” [20]). We made an effort to promote diversity and inclusion by covering...
participant travel expenses, refraining from jargon and acronyms, and enabling remote participation.

We set out to create a small team of intrinsically motivated innovators working voluntarily together to realize innovative ways to tackle the thorny problem of systemic improvement of chronic illness care [10]. We developed a social media presence; contacted people directly via phone, in person, or email; and invited those interested to an online community of innovators using a private social networking platform. We looked for people who had implemented creative workarounds to the systems’ barriers (in other words, who had hacked the health care system) and who were eager to collaborate with likeminded others. We responded quickly and substantively to potential solutions and connected together people working on similar problems. These efforts were augmented by webinars where stakeholders shared their perspectives and relevant work and other webinars where ideas from different disciplines were integrated into the overall design [17].

There are inherent power gradients in health care, and we managed these by explicitly acknowledging this dynamic, by privileging patient and family voices (eg, framing the context for all design meetings by having a patient or parent share their public narrative), and by active facilitation to ensure that patient and family voices were included.

**Idealized Design**

We used the idealized design process [21], a systematic process for creating and implementing new ideas through five steps (design, prototype, pilot, implementation, scale-up, and spread). Phase 0, the design phase, is an iterative process of observation, synthesis, and screening. It is focused on generating new ideas that could lead to a fundamental redesign to better meet the needs of users of the system. The 10-month iterative process consisted of interactions both synchronous (conference calls, webinars, and face-to-face meetings) and asynchronous (email and a private social media site). This process was punctuated by three design meetings, the objectives of which are provided in Textbox 1.

**Observation—Environmental Scan, Goal-Directed Design, and Needs of the Users**

Observation is the primary method for understanding patient needs and for generating ideas to meet those needs [12]. We used three techniques for observation—an environmental scan, goal-directed design, and understanding the needs of users. Innovations emerge from the inferences drawn from these observations.

We conducted an environmental scan to identify ideas and concepts that could fulfill user needs. We used a broad set of tools including key informant interviews, literature review, Internet searches, and group discussions.

We used goal-directed design—described in-depth elsewhere [16]—to understand human needs, as well as ideas that may satisfy those needs within a complex system. Goal-directed design begins with ethnographic and synthesis methods that generate personas—research-based composites of potential users of the new system—and scenarios that depict personas realizing their goals through interacting with the new system.

Understanding user needs and the current state informs the development of high level outcome measures. During design meeting 1, participants predicted the needs and goals of patients or families, clinicians, and researchers. These were synthesized to create measurement concepts and then outcome measures to assess the ability of a C3N to achieve its aims, centered on human needs, during the testing phase.

**Synthesizing—Concept Design, Key and Secondary Drivers, and Innovations**

Design synthesis is the abductive process of organizing and manipulating observations, data, and ideas into a coherent whole, both synthesizing observations into interventions and synthesizing interventions into a concept design. Observations were synthesized into a conceptual framework (a high-level description of what a C3N is and ought to do), a set of key drivers that must be in place to change the outcomes, and a set of intervention concepts, called secondary drivers, that might bring about these key drivers [22].

During design meeting 2, participants used the secondary drivers and personas to generate scenarios and a set of innovations (prototypes) that could be tested for their ability to change outcomes, whether individually or in combination. Participants also screened and elaborated on the KDD, metrics and targets, and possible innovations.

**Screening—Prioritizing Interventions and Assessing Coverage**

Ideas were screened through criteria such as “Is it...desirable?” “...different?” “…feasible?” and “Will the idea move us beyond current best practice?” In design meeting 3, participants rated each intervention concept based on potential impact and degree of understanding or knowledge for implementation. Using this 2 x 2 matrix, intervention concepts could be rated high impact and high knowledge (implemented relatively easily to good effect), high impact and low knowledge (could have a positive effect but require further development), or low effect with or without high knowledge (screened out because of little or no expected impact). Intervention concepts classified as high impact but low knowledge became prototyping candidates owing to their potential for teaching us the most about particular interventions employed as part of a peer production knowledge network such as C3N.
Textbox 1. Objectives for design meetings.

<table>
<thead>
<tr>
<th>Design meeting 1 objectives:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Collaborative chronic care network (C3N) design meeting participants:</td>
</tr>
<tr>
<td>• Meet and develop a level of comfort and familiarity with each other</td>
</tr>
<tr>
<td>• Develop an appreciation for the broad range of expertise, experience, and approaches that each brings to the design process</td>
</tr>
<tr>
<td>• Develop a shared understanding of the common purpose of the C3N</td>
</tr>
<tr>
<td>• Understand the phase 0 design process</td>
</tr>
<tr>
<td>• Develop a shared initial vision of the final C3N</td>
</tr>
<tr>
<td>• Understand how their work fits into the C3N</td>
</tr>
<tr>
<td>2. Develop a clear articulation of the problem(s) that the human-centered design should address</td>
</tr>
<tr>
<td>3. Understand the health care ecology model</td>
</tr>
<tr>
<td>4. Give feedback to a proposed set of specific characteristics (for patients, families, clinicians who are part of networks, and researchers) that will be sampled for during the human-centered design process</td>
</tr>
<tr>
<td>5. Develop predictions of what the end users will say about their needs</td>
</tr>
</tbody>
</table>

Design meeting 2 objectives:
1. Introduce and reintroduce C3N participants to each other
2. Obtain input on a refined vision, purpose, values and principles, and metrics
3. Screen and elaborate a proposed initial system driver diagram
4. Screen and elaborate information, technical and experience architectures for the patient-facing portions of the design
5. Identify possible studies and prototypes

Design meeting 3 objectives:
1. Align participants around the C3N system driver diagram
2. Rate the secondary drivers as to their contribution or importance to desired system outcomes
3. Ensure that potential prototypes or work products for the next phase of the project sufficiently cover the highly rated components of the system driver diagram
4. Prioritize the prototypes or work products for the next project phase based on those that will advance the improvement and research and development efforts
5. Begin to scope the required effort, team composition, and other resources for a number of highly rated prototypes or work products

Results

Participants
The design process began with a relatively small team of 25 members. By actively reaching out to more participants and inviting them to “make a difference” and “be part of the solution,” we increased the number of people involved over the design phase to 150 people (9 C3N team, 28 clinicians, 9 designers, 6 informatics experts, 11 patients, 5 parents, 54 collaborators, and 27 project and research staff), exchanging over 1700 posts and messages on the private social media site. There were 8 clinicians, 3 patients, 3 parents, 2 staff, and 18 C3N team members and collaborators at design meeting 1; 6 clinicians, 3 patients, 2 parents, 6 staff, and 18 C3N team members and collaborators at design meeting 2; and 7 clinicians, 2 patients, 0 parents, 3 staff, and 12 C3N team members and collaborators at design meeting 3. All work teams that formed had representation from patient or family, clinician, and researchers stakeholder groups.

Observation

Environmental Scan
Interviews with thought leaders provided the following C3N design imperatives:

• Design for all stakeholders at once—not separately for patients, clinicians, and researchers.
• Technology is only a means to an end. The focus of design must be on enabling people to gracefully achieve their goals, with appropriate technology deployed in service of those goals.
• Notwithstanding the above, an upgradeable set of modular technologies is likely to prove to support system evolution.
• Design with acknowledgement of health care as a service that is coproduced, not a product to be delivered. This shifts the paradigm from health care as a transaction to health care as shared work [23].
Design to enable large, diverse groups of people to identify and test many solutions to many problems. No one person and no one solution will transform care, outcomes, and cost.

C3N leaders must empower others to achieve common aims under uncertain conditions. This entails fostering a rapid learning ethos and an embrace of failure in service of learning.

Our environmental scan uncovered 64 people, organizations, products, or services that provided inspiration for parts of the C3N. These included 5 blogs, 10 information clearinghouses, 3 design firms, 4 potential funders, 8 experts or innovators or innovations, 15 networking or community platforms, 3 stakeholder representatives, 4 information technology innovations, and 12 thought leaders. Examples include commons-based peer production models (eg, Linux, Wikipedia, TripAdvisor, Slashdot, and Science Commons); patient communities (eg, PatientsLikeMe, Crohnology, CureTogether.com, and e-patients.net); crowd-sourcing platforms such as Innocentive and Many Eyes; and QI and research collaboratives such as the Northern New England Cardiovascular Disease Study Group and the Children’s Oncology Group.

We also recognized challenges. Unlike Wikipedia or TripAdvisor, medicine has inherent power and knowledge differentials and regulatory and oversight constraints. Developing the right mix of incentives to engender collaborative behavior is challenging, as is attracting individuals to contribute and foster contributions. A free-for-all where everyone’s opinion is equal risks introducing and propagating harmful ideas and suggestions.

Textbox 2. Purpose, vision, values, and principles for the collaborative chronic care network (C3N).

Creating a C3N

1. Purpose:
   - To enable patients and families, clinicians, and researchers to work together to create a Collaborative Chronic Care Network for inflammatory bowel disease (IBD) that transforms the outcomes and experience of illness and care, spawns innovations, and accelerates discovery and the application of new knowledge. Working with ImproveCareNow, we will design, create, and test new approaches to transforming the system of chronic illness care for IBD. By (date), the project will have produced working prototypes of components of the new system that can improve the outcomes, process, and experience of care and increase the production of innovations in care delivery and new knowledge.

2. Vision:
   - To be healthier together
   - Patients, families, clinicians, and researchers all have the same goal when it comes to chronic disease—for those affected to be healthy and live gracefully with a condition that they didn’t ask to have
   - The C3N will enable ImproveCareNow to become a collaborative innovation network—a community with shared purpose, values, tools, and technologies (both human and digital) to enable patients, families, clinicians, and researchers to share responsibility for achieving dramatically better health for children and adolescents with IBD

3. Values and principles (How we behave in the community and act toward one another):
   - Hope and compassion (to cause the enthusiasm, curiosity, and the will to solve problems)
   - Privacy (must be data literate to participate, patients own their data: have rights of possession, use, and disposal)
   - Trust (individuals demonstrate credibility, information is credible; scholarly norms for attribution; openness)
   - Shared responsibility for outcomes—we all have the responsibility to improve the health of the entire community (the entire population of patients with IBD)—you can make a difference, and you are expected to
   - Urgency and daring to create and try new ideas
   - Creativity and innovation (to generate and test new solutions)
   - Self-determination (agency)

4. Descriptions of keywords:
   - Collaborative—patients or families, clinicians, and researchers engaged as partners in a shared task.
   - Community—a distributed, voluntary organization that is interdependent, has shared responsibility, and is greater than the sum of its parts.
   - Shared purpose—improve the health of the entire IBD community
   - Common values—compassionate, safe, trust (privacy and credibility), open, self-determination, or agency
   - Flexible set of tools—human (quality improvement, leadership training, motivational interviewing, social networks, and incentives) and information (asynchronous communication, social media, network analysis, data mining, and multimedia) technologies
Goal-Directed Design

The goal-directed design method is presented elsewhere [16]. Personas representing all key stakeholders—patients, parents, physicians, nurses, and researchers—were created. Overall, the personas and scenarios enabled the design participants to maintain a focus on key users and how they might interact with the new system. The main contribution of this method was to keep the focus of the design on people—patients, parents, clinicians, and researchers—and on helping people meet their goals rather than to focus on the tasks required to meet these goals. For example, in one scenario, the patient persona, Bianca, is connected by her nurse, Vicki, to other patients on a virtual platform, where she overcomes her sense of isolation by sharing experiences with others similar to herself. In this case, the design was in response to Bianca’s goal to avoid isolation and remain connected to others and to Vicki’s goal of making sure patients have the support necessary to thrive. Although important tasks or features are implied in this scenario (eg, identity authentication, secure messaging, and community moderation), these were purposefully tabled to be addressed later in the design process.

High-Level Outcome Measures

Including all relevant stakeholders as codesigners enlarged the discussion of relevant outcomes beyond traditional clinical measures. By focusing on people and their goals or needs and by insisting that the system must meet the needs of all people, we were able to arrive at a set of measures that reflected the multistakeholder perspective. The following system performance measures were proposed and approved:

- Participation, engagement, and interaction among all types of users as measured by attendance at webinars, monthly calls, and community conferences, as well as contributions of data and ideas.
- Health outcomes (eg, steroid-free remission and improved quality of life) as measured by physician global assessment.

Textbox 3. Intervention concepts prioritized for further testing, based on high ratings on potential impact and low ratings on understanding or knowledge.

| Mentoring in inflammatory bowel disease (IBD) clinic |
| Leadership training |
| Privacy education |
| Facebook connector app and community building |
| Branding ImproveCareNow as a collaborative chronic care network (C3N) |
| Model care or quality improvement metric explorer |
| Android device—gateway to C3N |
| Virtual camp oasis |
| Self-management support curriculum |
| Open-source practice Wiki |
| Patient driven n=1 trial |
| Restructured IBD education day |
| Patient interface—virtual C3N |

Reliability and effectiveness of chronic illness care (eg, more appropriate medication use and disease activity monitoring) as measured by the degree to which a bundle of clinical interventions were delivered as part of clinical care.

- Self-management, as measured by self-report of adherence.
- Production of new knowledge and discoveries as measured by research products, including grants, abstracts, presentations, and publications.

Synthesis

Concept Design and Key Driver Diagram

Representatives from all stakeholder groups cowrote the purpose, vision, values, and principles for the C3N (Textbox 2) during design meeting 1, and the KDD (Figure 2) during design meeting 2. Taking a multi-stakeholder perspective forced the design team to consider the new system not as a system for doctors or a system for patients, but rather as a system for people. This, in turn, allowed ideas from outside of health care to be brought to bear in the concept development.

Generating Ideas and Scenarios

Design meeting 2 also resulted in ~140 potential innovation ideas. We deduplicated and combined the ideas into 33 unique potential innovations.

Priority Setting

During design meeting 3, a total of 20 intervention concepts were rated as having high impact and high understanding or knowledge. These were interventions that ImproveCareNow was either doing currently or else were sufficiently specified so that no further design or testing was necessary. These should simply be done. There were 13 intervention concepts rated as having high potential impact and low understanding or knowledge about how to implement. These intervention concepts, listed in Textbox 3, were selected for further development and testing.
Discussion

Principal Findings

Our codesign process resulted in a community of over 100 people willing and able to self-organize to pursue a shared overall concept design for the ImproveCareNow C3N, a logic model for bringing about this system, and 13 potential innovations likely to increase awareness and agency, make it easier to collect and share information, and to enhance collaboration. Developing and testing these potential innovations to determine the degree to which they could collectively bring about the C3N were the actionable next steps for ImproveCareNow.

It is not intuitive that thousands of people could self-organize and collaborate to achieve shared aims. But we see they do across many industries. In addition to Wikipedia and other examples uncovered in our environmental scan, more recent examples such as AirBnB, Uber, Lyft, and crowd funding sites such as Kickstarter and GoFundMe affect the lives of more and more people. The C3N design is a way to translate peer production to health care.

Distributed networks are especially relevant to children with chronic diseases that the NIH identifies as rare diseases [24] because no single health center has a sufficient number of patients to produce generalizable knowledge [25]. This state of affairs can result in a slow pace of knowledge acquisition and outcome improvement. Networks are also of growing importance to clinicians to support collaborative learning and application. Networks of patients and the rise of the e-patient movement (eg, Patients Like Me, Association of Cancer Online Resources, Crohnology, and Society for Participatory Medicine) have enabled patients to collect their own data for research and to support one another. But the potential of these networks to impact the overall chronic care system is limited because they operate in a siloed manner.

There is growing awareness that health care is a coproduced service—that professionals and patients create value through collaborative interactions [23]. Traditionally and appropriately, the focus has been on interactions within each clinical encounter [26]. By enlarging the focus, considering one-to-many and many-to-many interactions, and applying peer production principles, the C3N design recognizes the value of networks in health care. Fjeldstad and colleagues suggest collaborative networks share a common architecture, including actors who have the motivation and ability to self-organize; a commons where resources are created and shared; and structures, protocols, and processes that facilitate multi-actor collaboration [11]. C3N design intervention concepts can be viewed through this Actor-Oriented Architecture lens.

We conceived of the C3N as a health care system in which patients (and their families), providers (physicians and other clinicians), and researchers could collaborate, at large scale, to achieve shared aims. By codesigning with representatives of all stakeholder groups, we were able to translate this idea into a design concept, including a set of measures, a logic model, and a set of innovations that could be tested together to achieve the goal of improving care, spawning innovation, and accelerating research. The C3N model challenges the dominant chronic illness care paradigm that views patients as objects on which to intervene, structures care around episodic one-to-one patient-physician interactions, and assumes an inherent power differential based on knowledge. The C3N is designed to engage patients as coequals in care delivery, designing innovations, and research; make learning continuous; and level the knowledge gradient.

Challenges Encountered

We encountered several challenges during this design phase. Because a C3N had never been created before, we did not know what the end product ought to be, and this was frustrating to some stakeholders who wanted to know what the answer was. Over the course of the design process, most stakeholders came to realize that there was no predesigned product and that the point of the codesign process was to come up with this answer. Another challenge was managing expectations of how transformative the changes would be. Some stakeholders were nervous that the design would be too much of a change, whereas others feared the opposite. We regularly introduced the topic of change and attempted to calibrate expectations, in part, by reiterating that the codesign process itself would ensure that the final product was acceptable to the community. A third challenge was the need to translate across stakeholders so that a common perspective and even a common language emerged. Words like community and social network are used in common parlance but have specific scientific meanings that may be different from their connotations.

Limitations

Because a C3N had never been made before, there was no way to know in advance what steps to follow to bring it about. We developed, rather, a collaborative team of more than 100 diverse stakeholders aligned around a common goal and with a common plan for testing our way into this new system. This team was able to identify user needs and generate a sufficient set of novel ideas that could be potentially transformative for ImproveCareNow [21].

The clinicians and patients with whom we worked are likely to be systematically unrepresentative of the population, having relatively high levels of skills, insights, or resources. This set of conditions risks creation of a design that would work only for these users. Our rationale for this strategy is based on von Hippel’s theory of lead-users [27] that posits that in the case of a new product or product category, most users will not have the real-world experience necessary to contribute to its development. Lead users are those whose current strong needs will become general in the near future. They often attempt to fill their needs by creating novel solutions. Accordingly, we identified and worked with lead users in the codesign process. In addition, we guarded against a narrow design through the use of personas and scenarios developed through goal-directed design, which offers design targets more representative of potential users with fewer advantages.

The generalizability of the C3N design is unknown. Although this design was built for IBD, the noncategorical approach to
chronic illness care [28] and the chronic care model [26] both suggest that there are common problems faced by people and common processes necessary for good clinical care across conditions. This would argue for generalizability. However, not all chronic disease is like IBD: when patients are in remission, patients with IBD feel well and may forget about the disease. Other illnesses such as diabetes or cystic fibrosis require relentless attention to care. We intend to test the C3N design in other conditions.

Finally, the absence of formal feedback from our codesign participants limits our ability to understand how acceptable different users found the process.

Conclusions

Our current health care system cannot achieve the results we need. Incrementally improving the current system is not enough, but designing a new system is a daunting task. Our experience suggests that codesigning with representatives from all relevant stakeholders, using the idealized design process, can result in a potentially transformative design for the chronic care delivery system.

Acknowledgments

This work was funded by the NIH through the NIH Roadmap for Medical Research (R01 DK085713-01). The authors would like to acknowledge the ImproveCareNow Network for their contribution to this work.

Conflicts of Interest

None declared.

References

Supporting Accurate Interpretation of Self-Administered Medical Test Results for Mobile Health: Assessment of Design, Demographics, and Health Condition

Jess C Hohenstein1, MS; Eric PS Baumer2, PhD; Lindsay Reynolds1, PhD; Elizabeth L Murnane1, PhD; Dakota O'Dell3, PhD; Seoho Lee4, PhD; Shion Guha5, PhD; Yu Qi1, MPS; Erin Rieger6, BS; Geri Gay1, PhD

1Department of Information Science, Cornell University, Ithaca, NY, United States
2Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, United States
3School of Applied and Engineering Physics, Cornell University, Ithaca, NY, United States
4Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, United States
5Department of Mathematics, Statistics, and Computer Science, Marquette University, Milwaukee, WI, United States
6Department of Chemistry, Rice University, Houston, TX, United States

Corresponding Author:
Jess C Hohenstein, MS
Department of Information Science
Cornell University
107 Hoy Road
Ithaca, NY, 14853
United States
Phone: 1 607 254 5347
Email: jch378@cornell.edu

Abstract

Background: Technological advances in personal informatics allow people to track their own health in a variety of ways, representing a dramatic change in individuals’ control of their own wellness. However, research regarding patient interpretation of traditional medical tests highlights the risks in making complex medical data available to a general audience.

Objective: This study aimed to explore how people interpret medical test results, examined in the context of a mobile blood testing system developed to enable self-care and health management.

Methods: In a preliminary investigation and main study, we presented 27 and 303 adults, respectively, with hypothetical results from several blood tests via one of the several mobile interface designs: a number representing the raw measurement of the tested biomarker, natural language text indicating whether the biomarker’s level was low or high, or a one-dimensional chart illustrating this level along a low-healthy axis. We measured respondents’ correctness in evaluating these results and their confidence in their interpretations. Participants also told us about any follow-up actions they would take based on the result and how they envisioned, generally, using our proposed personal health system.

Results: We find that a majority of participants (242/328, 73.8%) were accurate in their interpretations of their diagnostic results. However, 135 of 328 participants (41.1%) expressed uncertainty and confusion about their ability to correctly interpret these results. We also find that demographics and interface design can impact interpretation accuracy, including false confidence, which we define as a respondent having above average confidence despite interpreting a result inaccurately. Specifically, participants who saw a natural language design were the least likely (421.47 times, P=.02) to exhibit false confidence, and women who saw a graph design were less likely (8.67 times, P=.04) to have false confidence. On the other hand, false confidence was more likely among participants who self-identified as Asian (25.30 times, P=.02), white (13.99 times, P=.01), and Hispanic (6.19 times, P=.04). Finally, with the natural language design, participants who were more educated were, for each one-unit increase in education level, more likely (3.06 times, P=.02) to have false confidence.

Conclusions: Our findings illustrate both promises and challenges of interpreting medical data outside of a clinical setting and suggest instances where personal informatics may be inappropriate. In surfacing these tensions, we outline concrete interface design strategies that are more sensitive to users’ capabilities and conditions.

http://humanfactors.jmir.org/2018/1/e9/
However, an important and understudied consideration is how individuals’ “lived informatics” practices [5,6] and how they one [1]. The research community has documented such weight, blood sugar, sleep, etc) for themselves or for a loved

10 US adults now track a health indicator (blood pressure, mood, and passively, a wide range of data that have traditionally been these personal devices allow individuals to track, both manually

and interpreted by individuals outside of traditional health care settings [1]. The widespread use of personal tools for collecting, analyzing, and providing feedback about health data poses broad questions regarding how people make sense of this information. What kinds of medical data are appropriate to self-monitor? Without relevant training and practice, can laypersons accurately interpret their own health measures? Furthermore, are people confident in their ability to take control of their own health in these ways, without consultation with a health care professional?

This paper explores these questions through both small-scale interviews (N=27) and a large-scale survey (N=303) that examine how various interface designs impact diverse users’ accuracy and confidence in interpreting the results of medical tests. In doing so, this paper makes several contributions:

1. A characterization of the advantages and challenges of using personal informatics technology to self-gather and interpret various types of medical data, including insights into situations where hesitation is warranted before deploying mobile health–based interventions

2. Definitions for measuring 2 specific problematic self-assessment scenarios, false confidence and false hesitance, along with our results regarding how various feedback formats and demographic attributes can predict these constructs

3. A set of concrete design recommendations to support users’ accuracy and confidence in interpreting feedback from mobile health tests

4. A general discussion of how future personal health systems can move in more tailored directions to support a greater harmony among specific interface components, user characteristics, and qualities of a monitored aspect of health

Health Apps and (Self-)Tracking

In the United States, ownership of mobile technology is incredibly pervasive, with 90% of people owning cellphones and 64% of people owning smartphones specifically [2,3]. Globally, it is estimated that by 2020, 80% of adults will have a smartphone [4]. The extensive data-capture capabilities of these personal devices allow individuals to track, both manually and passively, a wide range of data that have traditionally been gathered in a clinical or laboratory setting. For instance, 7 in 10 US adults now track a health indicator (blood pressure, mood, weight, blood sugar, sleep, etc) for themselves or for a loved one [1]. The research community has documented such individuals’ “lived informatics” practices [5,6] and how they collect and use personal data to make changes in their lives [7].

However, an important and understudied consideration is how people are interpreting these self-gathered results, including the accuracy of their interpretations.

Patient Interpretation of Medical Test Results

According to fuzzy trace theory, when making decisions, people rely on the “gist” of the information they receive, or their interpretation of the bottom-line meaning, instead of verbatim details, which explains why precise information is not necessarily effective in supporting medical decision making [8]. Previous work investigating the effects of patients being given direct access to their (clinician-gathered) personal health records has shown mixed results [9]. Specifically, although patients can feel an enhanced sense of control over their health, direct data access brings risks, including patients incorrectly interpreting the data or taking the wrong action in response.

Another concern of direct medical data access relates to issues of health literacy. A trained clinician can interpret test results with an implicit awareness of how values map onto severity or where thresholds for action lie—information that is unfamiliar or invisible to most patients [10]. Such challenges are compounded by the fact that in the United States, low numeracy is widespread, and written information about tests and their results are often provided at higher reading levels than many patients can manage [11] or in presentation formats that are perceived as uninformative [12]. Similarly, studies have found that many people experience difficulty in interpreting health information from graphs (ie, low graph literacy) [13,14]. Relatedly, diverse groups of people may respond differently to the same image-based feedback because of individual differences (eg, gender [15]) in visual perception, such as processing static versus animated images [16] or in the strength of reactions to pleasant or unpleasant imagery [17-19].

Furthermore, some groups of patients have highly variable relationships with health care as a whole. Racial and ethnic disparities in medical access and quality have been extensively documented [20], and some groups are more likely to experience bias and a lack of cultural understanding in health care [21]. Such problems could potentially translate into less involvement in the self-monitoring process to begin with or less confidence in interpreting health data. On the other hand, patients with higher levels of education may be more self-monitoring savvy and confident, given that research finds they are often better able to manage self-care regimens [22], are faster to adopt new medical technologies [23], and are more likely to use preventative care [24].

Altogether, such differences in comprehension and confidence between groups must be considered in the context of personal informatics and the interpretation of health data. Given the widespread acceptance of smartphones and self-tracking technologies, sophisticated personal medical tests will be a reality for the general population in the near future. The important implications of these tests require informed design of the interfaces used to present test results for general use.
Although significant prior work has focused on individuals’ use of personal informatics tools [25,26], there is a lack of research that considers how various design strategies might impact users’ ability to interpret their own health measures outside of a clinical setting, along with their confidence in these interpretations.

Methods

Overview

This paper investigates individuals’ interpretation of health data outside of a clinical context. To do so, we used NutriPhone [27], our prototype system (see Multimedia Appendix 1) that transforms any mobile device into a point-of-care biomarker assessment tool by combining blood testing strips, a custom hardware accessory, image analysis software, and a user-facing app that delivers diagnostic reports.

Preliminary Investigation

To gain qualitative insight into how people interpret medical data through NutriPhone, our preliminary investigation (Cornell Institutional Review Board Protocol ID#1410005065) used direct observation and dialogue in an interview-based lab study. We used an on-campus recruiting system to recruit participants (N=27, 20 female, aged 18-45 years). A total of 24 were undergraduate students who were compensated with course credit, and the remaining 3 were academic staff who volunteered their time. Interviews lasted approximately 10 min, were conducted in person, and were audio-recorded and transcribed.

Because the goal of the preliminary investigation was not to identify which design elements maximized interpretation accuracy but rather was aimed at observing and discussing participants’ process of interpretation and how various design choices impact it, we used a hybrid interface design. Specifically, we combined textual, graphical, numerical, and color components to reflect the predominant formats of visual feedback used by personal informatics systems and to appeal to multiple types of literacy [28]. For the health indicator, we chose to present vitamin B12 levels. Because vitamin B12 deficiency is fairly uncommon in developed nations, doctors rarely test vitamin B12 in isolation or discuss it with their patients [29], meaning our participants were unlikely to have prior knowledge about and would need to rely on our interface for interpreting the data. We implemented 2 versions of the interface, which can be seen in Figure 1: the “Healthy Result” (left) displayed a B12 level within the US National Institutes of Health–recommended reference range, and the “Low Result” (right) displayed a B12 level lower than this reference range [30].

To begin, participants were provided with a link to access the NutriPhone app on their personal smartphones, with the interface variant randomly assigned (14 and 13 participants saw the healthy and low variants, respectively). We told participants that the purpose of the app was to “help people run blood tests on their own without a health care practitioner,” and they were asked to imagine that they had already completed the testing procedure. We next asked participants to describe their test result and then followed up with questions about how they understood (or did not understand) their result, their usual method for interpreting medical test results, and overall impressions.

The results from the preliminary investigation pointed in 2 directions. First, 25 out of 27 participants (93%) correctly interpreted the test results (ie, correctly answered that their result was high or low when viewing with the high or low interface). However, despite their overall accuracy, 20 out of 27 participants (74%) also expressed confusion and doubted their interpretations. When asked what their test result meant, responses were often a variant of “I don’t know” or “I have no idea what that means.” Such doubt is important to consider, as it can inhibit the translation from insight to action, even if a person’s interpretation is in fact accurate.

Figure 1. NutriPhone interfaces presented to participants in the preliminary investigation. Two variants of the interface showed a healthy result (left) and an unhealthy result (right). Both variants incorporated textual, graphical, numerical, and color design components.
Participants also wondered about the potential influence of individual or demographic characteristics on their result. Finally, several participants wanted to see an average or “typical” score or range to help them situate their results within the larger population, whereas a few wanted a more binary presentation that simply indicated whether their result was problematic or not.

Our preliminary findings left us with some unexpected outcomes and unresolved questions. In particular, we did not anticipate that such highly accurate interpretations would be accompanied by such confusion and uncertainty. Furthermore, we did not directly analyze which interface components supported correctness or contributed to doubt. Finally, having only looked at 1 biomarker, we were left wondering how participants would interpret other more well-known biomarkers with different reference ranges and whether those conclusions would be made more confidently.

**Main Study**

To pursue these goals, our main study (Cornell Institutional Review Board Protocol ID#1410005065) focused on 3 different interface designs that present numerical, textual, and graph-based feedback:

1. **Number**: Biomarker level is presented as a number, providing only the raw measurement that would result from a blood test.
2. **Natural language**: Biomarker level is presented using natural language text that explains whether the biomarker level is considered low or high.
3. **Graph**: Biomarker level is presented graphically, with a marker at the measured value. The one-dimensional chart includes “low” or “high” anchors to provide orientation.

As mentioned earlier, these design styles were chosen to reflect the conventional feedback formats found in personal informatics systems and to appeal to distinct types of literacy [28]. Although informal pilot testing indicated that participants typically correctly interpreted green as healthy and red as unhealthy, we chose not to test a color-based feedback design because of inherent accessibility issues. Specifically, other cultures may ascribe these colors with different meanings [31], and the widely used “stoplight”-style color system for risk presentation [32] is indistinguishable for individuals with deuteranopia (insensitivity to green light, commonly known as red-green colorblindness).

Next, to broaden our variety of examined medical data, we focused on the following 3 biomarkers: vitamin B12, procalcitonin (PCT), and cholesterol. First, these biomarkers vary in terms of participants’ expected prior familiarity with them. Similar to B12, participants were unlikely to have prior knowledge of and know how to interpret PCT, which is used to diagnose bacteremia and septicemia [33]. In contrast, cholesterol is a more commonly known health marker, making participants more likely to be aware of what constitutes healthy levels. Our selected biomarkers also vary in terms of whether a higher or lower measure constitutes a healthier or an unhealthier result. As previously discussed, health consequences effectively only exist for low levels of vitamin B12. Conversely, PCT is problematic at high levels and has no medical consequences for very low or zero levels, and cholesterol similarly carries medical risk only at high levels.

For each of the 3 designs, we created mock-ups for each of the 3 biomarkers, resulting in 9 interface variants, as seen in Figure 2. Each variant included a “healthy reference range” for the respective biomarker at the bottom of the result screen. These reference ranges resemble what a patient would receive in a clinical setting, and participants’ comments from the preliminary investigation suggested that these ranges would facilitate interpretation. Because our preliminary findings showed no statistical difference between accurate interpretation of healthy or unhealthy results, we chose to display only unhealthy results.

**Participants**

To examine individuals’ interpretation, confidence, and overall reaction to these various interfaces, we deployed a Web-based survey in September 2016 through Qualtrics, a system through which we enlisted 303 participants (155 female, 147 male, 1 bigender), who received various incentives (cash, airline miles, redeemable points, etc) for their participation. After providing Qualtrics with the survey questions and format along with the number and desired demographics of participants, they performed the process of carrying out the survey. We excluded 2 respondents from our analysis: 1 bigender respondent, both to prevent undue influence on the results and to prevent potential deanonimization, and 1 respondent who entered an age of 6 years, which we considered as a typing error considering Qualtrics only recruits adults. This left 301 participants for the main analysis.

Demographic screening criteria based on Pew’s omnibus Internet survey [34] were used to ensure a diverse, demographically representative sample of US Internet users. Ages ranged from 18 to 90 years (mean 45.96, median 45, SD 16.34). Of 301 respondents, 100 had a 4-year degree (33.2%), 67 had some college degree (22.3%), 48 had a high school degree (15.9%), and 45 had a professional degree (15.0%). Annual household incomes ranged from US $40,000 to more than US $200,000 (mean US $88,210, median US $80,000, SD US $3142). Racially, 201 out of 301 respondents identified as white, 66.8%; 59 identified as Hispanic, Latino, or Spanish origin (19.6%); 37 identified as black, African American, or Negro (12.3%); and 12 identified as Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, or other Asian (4.0%). Racial categories were not mutually exclusive; individuals who identify as multiracial were allowed to select multiple races.

**Procedure**

Participants first gave informed consent after reading about the purpose, time commitment, question types, risks and benefits, confidentiality, data storage, and principal investigator for the study (see Multimedia Appendix 2). Next, participants were given background information about NutriPhone and then randomly assigned to 1 of the 3 biomarkers (B12, PCT, or cholesterol). Adapting materials from Mayo Clinic [35] and Medline Plus [21,36], we then gave participants some background about that biomarker, including medical consequences of and how to counteract unhealthy levels.
Figure 2. Our main study tested 3 feedback designs (number, natural language, and graph—left to right) with 3 biomarkers (vitamin B12, procalcitonin, and cholesterol—top to bottom).

Next, they were shown an unhealthy result via 1 of the 3 interface designs (number, natural language, or graph) and asked a series of questions, starting with “My levels of [biomarker] are...” with choices of “too low,” “healthy,” “too high,” and “unsure.” We also asked participants about how confident they were in this interpretation. Participants were next asked a series of multiple-choice questions about how they might use NutriPhone, free-response questions about their general
impressions of the system, and demographic questions. A few attention check questions (e.g., “What planet are humans from?”) were deployed throughout the survey, and all participants correctly responded to these checks. Respondents were not able to change their responses after they had been submitted.

**Results**

**Operationalizing and Predicting Problematic Interpretations**

In undertaking our quantitative analysis of survey responses, we identified 2 problematic scenarios. We term the first scenario “false confidence,” which represents a respondent having above-average confidence despite inaccurately interpreting a result.

This situation is particularly concerning, given that it equates to a person incorrectly believing he or she is healthy when that is not the case. Second, we observed instances of what we call “false hesitance,” where participants accurately interpreted the result they saw but had below-average confidence. Although potentially less health hazardous than false confidence, such situations emerged as a consistent theme in our preliminary investigation and could still lead to hesitation or failure to take an appropriate course of action to address an unhealthy result.

To operationalize false confidence, we created a binary variable capturing both whether a respondent supplied an incorrect interpretation of the result and whether his or her confidence was above the mean confidence of all respondents who supplied incorrect interpretations. This approach labeled 33 out of 301 respondents (10.7%) with false confidence. Operationalizing false hesitance followed a similar procedure in which we identified respondents who interpreted the result correctly but with confidence below the mean confidence of other correct respondents. This approach labeled 66 out of 301 respondents (21.9%) with false hesitance.

To determine the factors most strongly associated with false confidence and with false hesitance, we constructed 2 binary logistic regression models, 1 for each outcome, using all subsets model selection. Potential predictors included experimental condition (i.e., the design variant and the health condition the respondent saw), age, gender, education level, household income, and 1 binary variable each for the racial categories of Asian, black, Hispanic, and white, as well as interactions between the design variant and each of the other potential predictors. Model selection for false hesitance failed to converge on a significant model. That is, no subset of the variables we collected significantly predicted which participants would correctly interpret the interface and yet have low confidence in their answer.

We therefore focus on false confidence, for which model selection resulted in a model with a $P$ value of .004, an area under the curve score of 0.77, and a McFadden pseudo $R^2$ of .15, all of which indicate a good fit. Table 1 presents the model’s details. Results are presented in terms of odds ratios; an odds ratio of $+2.0$ means that a 1-unit increase in that predictor equates to a participant being 2 times more likely to exhibit false confidence, whereas an odds ratio of $-2.0$ means that a 1-unit increase in that predictor equates to a participant being 2 times less likely to exhibit false confidence.

As Table 1 shows, we found several main effects and a few interaction effects, with the strongest significant effects relating to design, gender, race, and education. Specifically, participants who saw the natural language design were the least likely to exhibit false confidence by far, and women who saw the graph design were over 8 times less likely to have false confidence. The lack of false confidence shown for the natural language design makes sense given the fact that participants were screened for English-language proficiency, especially when compared with the graph and number designs, for which we would not expect to see false confidence given the widespread low graphical and numerical literacy in the United States [13,14].

Table 1. Model for false confidence, showing odds ratios and $P$ values. Main effects occur for the natural language design and for race. Interactions occur between the graph design and gender, the natural language design and race, and the natural language design and education.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Odds ratio</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graph design</td>
<td>+2.29</td>
<td>.56</td>
</tr>
<tr>
<td>Natural language design</td>
<td>−421.47</td>
<td>.02</td>
</tr>
<tr>
<td>Female</td>
<td>+1.06</td>
<td>.93</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>+25.30</td>
<td>.02</td>
</tr>
<tr>
<td>Black</td>
<td>+7.38</td>
<td>.19</td>
</tr>
<tr>
<td>Hispanic</td>
<td>+6.19</td>
<td>.04</td>
</tr>
<tr>
<td>White</td>
<td>+13.99</td>
<td>.01</td>
</tr>
<tr>
<td>Education</td>
<td>−1.14</td>
<td>.53</td>
</tr>
<tr>
<td>Graph design × female</td>
<td>−8.67</td>
<td>.04</td>
</tr>
<tr>
<td>Natural language design × female</td>
<td>−3.76</td>
<td>.27</td>
</tr>
<tr>
<td>Graph design × education</td>
<td>−1.07</td>
<td>.83</td>
</tr>
<tr>
<td>Natural language design × education</td>
<td>+3.06</td>
<td>.02</td>
</tr>
</tbody>
</table>
On the other hand, false confidence was more likely among participants who self-identified as Asian, white, and Hispanic. Finally, with the natural language design, participants who were more educated were, for each 1-unit increase in education level, approximately 3 times more likely to have false confidence. The observed findings regarding false confidence may reflect the findings of the National Assessment of Adult Literacy [37], which showed that white and Asian/Pacific Islander adults had higher average health literacy than adults of other races and that average health literacy increased with each higher level of educational attainment. However, our work is also somewhat at odds with these findings, as these groups were falsely confident in incorrect interpretations of health data, suggesting the possibility that groups with relatively higher health literacy could be more prone to unknowingly misinterpreting health data.

**Promises and Risks of Interpreting Self-Gathered Medical Data**

Encouragingly, 242 out of 328 participants (73.8%) interpreted their result accurately, in spite of the typical lack of common knowledge about vitamin B12 and PCT described earlier as well as the aforementioned low numeracy and graphical literacy rates [28,38]. These findings demonstrate the possibility of understandably conveying health information (even about less familiar health indicators) through mobile interfaces, as long as the design of that feedback provides enough context.

We saw additional glimpses into the potential of giving people access to their medical data as a number of respondents described a desire to use our prototype system to monitor their personal health. For example, participants expressed that the tool would be helpful for self-screening or could help ease nerves surrounding a health condition of personal concern. One participant told us how she could “save money from having to go to the lab to have my blood tested; I could do this all in my own home.” Other respondents envisioned using the system as a way to get actionable guidance when making health-related lifestyle changes and were open to the system additionally providing more prescriptive behavioral feedback, with 1 participant expressing that it would “...help control their health and be on top of things.”

At the same time, our results also highlight disadvantages of allowing individuals to interpret their own medical data and areas where personal informatics may be less appropriate. Although the majority of our participants correctly interpreted their results, many expressed confusion and questioned their interpretations: 20 out of 27 (74%) participants in the preliminary investigation expressed self-doubt, and 115 out of 301 respondents (38.2%) expressed low confidence in the main study. We believe that our addition of a “healthy reference range” with the result is largely responsible for this decrease in confusion between our preliminary and main studies. It is also possible that the main study’s survey-based methodology was more susceptible to social desirability bias compared with the more personal nature of the in-lab study, which may have encouraged participants to open up about interpretive insecurities.

**Discussion**

**Principal Findings**

The results of our study provide a mix of implications regarding whether or not (and if so, how) personal informatics tools should support individuals in gathering and interpreting their own medical data. Overall, we find that a thorough understanding of the target audience is necessary before deploying any personal informatics tool and, especially for tests with vital consequences [39]. Regardless, as self-tracking gains increasingly mainstream popularity, direct access to medical data is becoming a reality, making investigations into effective ways to communicate mobile health data imperative to the future of personal informatics.

**Design Constraints and Recommendations**

Overall, our findings suggest several design strategies for presenting mobile health data to maximize users’ ability to correctly and confidently understand them. Primarily, there is value in using a hybrid feedback design that includes multiple representational modalities (eg, numbers, words, visual graphics), as such a design allows a designer to tap into different literacies to increase a display’s effectiveness. We saw more accurate interpretation of results in our preliminary investigation, where we used a hybrid design, than in the main study, where participants viewed designs with only a single representational format. In cases where it is not possible to include multiple types of feedback in the interface (eg, mobile apps where screen space is limited or when presenting results from multiple tests simultaneously), we recommend ensuring that a design integrates text-based feedback, where natural language is used to convey whether a result is “healthy,” “high,” or “low.” The natural language design was least likely to cause false confidence among our participants, and among all of our tested.
design variants, the natural language and hybrid designs were interpreted correctly most often.

Next, we recommend including a "healthy reference range" along with any results given, especially for biomarkers or other health indicators with which a user is expected to have less preexisting knowledge. Participants’ confusion with the lack of a reference range in the preliminary investigation seemed to be alleviated once it was included in the main study.

Finally, it seems worthwhile to allow users to input personal details. Many of our participants expressed uncertainty about how such factors might influence their test results, which would in turn contribute to their lack of confidence in both their results’ reliability as well as their own assessment. In our main study, we captured and analyzed demographic variables such as age and gender, but participants also indicated their receptivity to supplying other personal data that can influence a given health condition (eg, a cholesterol diagnostic tool requesting weight information). Even for tests where these variables are not in fact relevant, such as for vitamin B12, the ability to input this information may alleviate the user concerns we observed and in turn increase their trust and acceptance of the system.

Furthermore, our findings about how individual differences can impact interpretation outcomes suggest that there is an opportunity to dynamically adjust an interface’s feedback format to use the representation least likely to cause confusion or misinterpretation for a given person.

Implications for Personal Health Informatics

The level of confusion and inaccurate interpretation observed in our investigation suggests situations in which personal informatics may be inappropriate. Our studies tested 3 health conditions and found that although the majority of participants could correctly interpret the data, their analysis was consistently couched in confusion. We also saw that different groups of people vary in their interpretation confidence and accuracy. These findings indicate that before a mobile health system is introduced, developers should first ensure that the biomarker being tested is one that users are comfortable self-tracking and produces results that people confidently understand how to appropriately act on. The ramifications of some users inevitably interpreting results incorrectly must also be considered, with situations in which a serious health issue goes untreated (ie, false confidence or inaccurate interpretation) being the most problematic.

With less well-known biomarkers or for populations who are more susceptible to making misinterpretations, we recommend using systems such as NutriPhone as a mediator between patients and health care providers. For example, a user could complete a routine blood screening using such a tool in the hours before an appointment with their clinician. Immediately after the test, the results are available for the patient, but the results are also sent to the clinician, who would discuss them during the appointment, including an interpretation of any abnormal findings and agreeing on a treatment plan together with the patient. If follow-up tests are appropriate, use of the tool could be continued for at-home monitoring. This scenario preserves many of the promises of personal health tracking while mitigating the potential risks our study identified. Patients would be able to perform ecologically valid self-tracking, interact directly with their medical data, and become empowered with a more active role and informed voice in their treatment. In addition, oversight by a health care practitioner would ensure appropriateness of follow-up actions, reduction of patient confusion, and avoidance of the aforementioned dangerous scenarios. Leveraging personal informatics technologies to transfer this type of health care management more directly into the hands of patients is attractive from an institutional perspective (eg, appealing to clinicians and insurance companies), especially in light of anticipated physician shortages in the United States [40], and our study indicates that patients themselves are receptive on a personal level as well.

Limitations and Future Work

Finally, we would like to point out potential limitations of our research and lay out room for future work. First, the results we presented to participants were pregenerated data, not actual outcomes. Displaying mock data is a common practice in system evaluation and still enabled us to gain insights into our key research questions regarding how people interpret medical data using a mobile health system outside of a clinical setting. Using mock data also imposed much less burden and privacy risk for participants, as they did not need to collect and share potentially sensitive health information. That being said, participants may react differently if interpreting real diagnostics about their actual health, especially considering personal medical data have been shown to carry strong emotional connotations [21], which we did not observe in this study. A natural future step is therefore to explore individuals’ interpretation of their own diagnostic results presented through a mobile health system.

Next, although the design elements we tested (numbers, graphs, and words) demonstrated significant differences, this study represents a partial exploration of a vast design space. Future work would do well to consider other elements that might appeal to different kinds of literacies [12,14,28,38] (eg, other types of visual charts or perhaps entirely different interaction modalities such as audio- or tactile-based feedback). Similarly, it would be desirable to expand investigations into additional types of medical results. For instance, data such as body mass index (BMI) could be especially valuable, given that for measures such as BMI, knowing simply whether or not one’s value is “within normal limits” may not be sufficient.

Finally, this study captured participants’ interpretations at one point in time. Previous research [41] has found that some patients feel that long-term self-tracking is “effortful and time-consuming” and sometimes give up the practice out of frustration. Future work would benefit from considering potential learning or habituation effects and emotions arising from viewing subsequent tests over an extended period of time, especially considering this would be a typical experience for an individual managing a chronic condition.

Conclusions

Medical technology is changing rapidly, with numerous devices and systems placing health information directly in the hands of patients. Personal mobile health tools that present feedback using formats similar to those we have examined in this research

http://humanfactors.jmir.org/2018/1/e9/
will likely become a similarly substantial part of medical care. With that day fast approaching, researchers and practitioners must be prepared to design effective tools that are not only comprehensible but also allow patients to be correct and confident in their interpretations and follow-up actions. For example, we find that user understanding is cultivated by natural language–based feedback as well as hybrid designs that integrate multiple different representational formats. Such design strategies and the broader implications identified by studies such as ours are key to ensuring that future generations of systems are appropriate and useable in nonclinical settings.

Acknowledgments
Funding for this research was provided by the National Science Foundation under award number 1430092.

Conflicts of Interest
None declared.

Multimedia Appendix 1
The NutriPhone system for point-of-care diagnosis and health monitoring consists of 3 parts: (1) a disposable test strip for blood sample analysis, (2) a portable optical reader that images the test strip, and (3) a platform-agnostic software app to process the images and provide a diagnostic result to the end user.

Multimedia Appendix 2
The preface text shown to participants gives background information about NutriPhone and, depending on the condition, background about a biomarker adapted from Mayo Clinic and Medline Plus.

References


Abbreviations

BMI: body mass index
PCT: procalcitonin

©Jess C Hohenstein, Eric PS Baumer, Lindsay Reynolds, Elizabeth L Murnane, Dakota O'Dell, Seoho Lee, Shion Guha, Yu Qi, Erin Rieger, Geri Gay. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 28.02.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
General Practitioners’ Attitudes Toward a Web-Based Mental Health Service for Adolescents: Implications for Service Design and Delivery

Mirjana Subotic-Kerry¹, PhD; Catherine King¹,², BLibSt (Hons), MClinPsych; Kathleen O’Moore¹, MSc, DClin (Psych); Melinda Achilles¹, BBPsy; Bridianne O’Dea¹,², PhD

¹Black Dog Institute, Randwick, Sydney, Australia
²Faculty of Medicine, University of New South Wales, Sydney, Australia

Corresponding Author:
Mirjana Subotic-Kerry, PhD
Black Dog Institute
Hospital Road
Prince of Wales Hospital
Randwick, Sydney, 2031
Australia
Phone: 61 2 9382 8509
Email: m.subotic-kerry@blackdog.org.au

Abstract

Background: Anxiety disorders and depression are prevalent among youth. General practitioners (GPs) are often the first point of professional contact for treating health problems in young people. A Web-based mental health service delivered in partnership with schools may facilitate increased access to psychological care among adolescents. However, for such a model to be implemented successfully, GPs’ views need to be measured.

Objective: This study aimed to examine the needs and attitudes of GPs toward a Web-based mental health service for adolescents, and to identify the factors that may affect the provision of this type of service and likelihood of integration. Findings will inform the content and overall service design.

Methods: GPs were interviewed individually about the proposed Web-based service. Qualitative analysis of transcripts was performed using thematic coding. A short follow-up questionnaire was delivered to assess background characteristics, level of acceptability, and likelihood of integration of the Web-based mental health service.

Results: A total of 13 GPs participated in the interview and 11 completed a follow-up online questionnaire. Findings suggest strong support for the proposed Web-based mental health service. A wide range of factors were found to influence the likelihood of GPs integrating a Web-based service into their clinical practice. Coordinated collaboration with parents, students, school counselors, and other mental health care professionals were considered important by nearly all GPs. Confidence in Web-based care, noncompliance of adolescents and GPs, accessibility, privacy, and confidentiality were identified as potential barriers to adopting the proposed Web-based service.

Conclusions: GPs were open to a proposed Web-based service for the monitoring and management of anxiety and depression in adolescents, provided that a collaborative approach to care is used, the feedback regarding the client is clear, and privacy and security provisions are assured.

(JMIR Hum Factors 2018;5(1):e12) doi:10.2196/humanfactors.8913

KEYWORDS
anxiety; depression; adolescent; general practitioners; internet
Introduction

Background

Anxiety and depression are prevalent among adolescents. These disorders are associated with significant disability, including educational failure, poor relationships, and impaired daily functioning [1-3]. If left untreated, adolescents with these issues are more likely to experience chronic mental health problems in adulthood, alongside reduced workforce productivity, lower income, and poorer quality of life [4,5]. Seeking professional help is crucial for early detection, prevention, and treatment [6]. In Australia, general practitioners (GPs) provide the most accessible primary health care for young people [7,8], and are the main clinical services used by adolescents with a mental illness and emotional or behavioral problem [9].

In Australia, general practitioners (GPs) provide the most accessible primary health care for young people [7,8], and are the main clinical services used by adolescents with a mental illness and emotional or behavioral problem [9]. GPs act as a first point of contact in the assessment and management of mental health issues and are well placed to identify mental health issues when youth present for other matters. GPs are a gateway to other services within the health system and facilitate access to rebated psychological therapy and psychiatric assessments. However, GPs face several barriers when treating youth mental health issues. These include inadequate training in the identification and management of adolescent mental health, a lack of confidence in recognizing youth mental health problems, and their consultation skills [10,11]. There is also a need for longer sessions, which are poorly remunerated [10,11].

Additional barriers relating to time constraints and poor linkages with other relevant services have also been reported, particularly among those in rural areas [12,13]. As such, many GPs feel ill-equipped and under-resourced when providing adolescent mental health care.

Web-based interventions, also known as e-mental health, have emerged as a safe, therapeutically effective, and acceptable referral option for common mental health concerns [14]. The advantages of using internet technology in the treatment of youth mental health have been well documented [1,15,16]. Web-based interventions are cost-effective, supplement standard therapy [17-20], and allow for greater dissemination of psychological treatments as access to the internet bypasses geographical distance, costs, and stigma [1,15,16]. Web-based interventions also improve the flexibility of psychological treatment, as they are available at a time suitable to the person and provide a sense of anonymity [21]. Two recent meta-analyses indicate that Web-based interventions are effective in reducing young people’s symptoms of anxiety and depression, and suggest that Web-based prevention and treatments are a viable alternative to face-to-face care [22,23]. In Australia, this has led to a national training initiative aimed at increasing GPs’ referrals to e-mental health. However, many GPs are still concerned about the use of eHealth. GPs cite complex systems, access issues, and privacy concerns as barriers to implementation in their practice [24,25]. Others have highlighted the lack of evidence and quality control for interventions [26], and the lack of time, technical skills, and financial compensations to implement eHealth properly [27].

In addition, youth living in rural and remote Australia may have limited access to the internet [28,29]. To address these barriers, the Black Dog Institute has designed and developed a Web-based mental health service, Smooth Sailing, to identify and treat anxiety and depression in secondary school students.

Delivered in the classroom, the Smooth Sailing service is based on the principles of stepped care [30], where the intensity of interventions is matched to individuals’ symptom severity. This service utilizes the internet to register students to a service where their mental health is screened. Using clinically validated algorithms, youth are automatically allocated to one of the four sequential treatment steps based on their symptom severity. Treatment intensity is matched to the step, and ranges from self-directed online psychoeducation and computerized cognitive behavioral therapy to individual face-to-face psychotherapy. The service has an e-monitoring component, in which students are sent mobile phone or email messages to assess their recent mood. In the school context, students who are at risk for suicidality or have severe levels of symptoms are automatically alerted to the school counselor who initiates consultation within 24 to 48 hours. The school counselor is then responsible for assessing risk, triaging, and referring to external support such as a GP, who can facilitate access to rebated psychological services and medication. GPs would then have access to the monitoring and alerting features. However, GPs’ attitudes toward this type of service model are unknown.

Acceptability research in health services aims to understand the extent to which people delivering or receiving an intervention consider it to be appropriate and suited to their needs, based on either anticipated or experimentation responses [31]. A study of the acceptability of the Smooth Sailing service among school counselors [32] found that personal beliefs, knowledge of e-mental health, internet accessibility, privacy, and confidentiality issues influenced their likelihood of use. Previous research on the adoption of eHealth among GPs has mainly been conducted in relation to adult health services [27]. In general, GPs and adult patients have expressed a positive attitude toward eHealth in adult health services, although GPs have also identified barriers such as time constraints, lack of skills, and lack of financial incentives to implement new ways of working [27]. In contrast, the adoption of new technologies in health care by youth mental health clinicians was less positive. Resistance to technology was based on a preference for face-to-face engagement with young consumers and a belief that the integration of new technology would create extra work in an already under-resourced environment [33]. Additional barriers identified by youth workers were related to skills, training, and concerns around confidentiality and the legal implications of online technology [33]. It is unknown whether GPs working with youth share these same concerns. Understanding the needs of GPs and their attitudes toward a Web-based, stepped care, youth mental health service is vital for designing and delivering high-quality mental health services.

Study Aims

This study aimed to examine GPs’ attitudes toward Smooth Sailing—a Web-based, stepped care, mental health service for reducing depression and anxiety among youth. This study will help to identify the factors that may facilitate greater uptake and effectiveness of the proposed service model among GPs.
and will help researchers and developers to better understand how e-mental health can be adapted for general practice.

**Methods**

**Study Design**
This is a multi-methods study consisting of interviews and an online survey. Ethics approval was obtained from the University of New South Wales (HC154456).

**Participants, Procedure, and Recruitment**
Participants were aged over 18 years, fluent in English, and currently working as a GP in Australia. Recruitment took place between May and June 2016 via the Black Dog Institute website, social media channels, word of mouth, and an electronic direct mail-out to professional networks. The study advertisements invited GPs who were interested in the mental health of young people, the acceptability of providing mental health care to adolescents via the internet, and those wishing to inform the development of an online service. Interested participants contacted the research team and were subsequently provided with a consent form via email. Interview details were arranged via email. All interviews were conducted over the phone and were audiotaped. Interviews were conducted by one of the researchers (KOM) and were approximately 25 minutes in duration. Before the interview, a detailed overview of Smooth Sailing was emailed to participants. The interview also began with a verbal description of the service model and an opportunity to discuss and clarify any aspects of the model. The online survey link was sent to all GPs at the completion of their telephone interview. Participants were reimbursed with an Aus $20 gift voucher. Recruitment ceased when saturation was reached.

**Interview Schedule**
The semistructured interview consisted of the following 4 questions:

1. What do you think of the proposed online stepped care model?
2. What do you think about the role of the GP in this model?
3. Do you have any concerns with this model and suggestions on how we could manage these?
4. What sorts of things would you need to integrate this model into your general practice?

**Survey**
The online survey consisted of 13 questions, with 6 questions assessing participant demographics and background factors including age, gender, state (New South Wales, Australian Capital Territory, Queensland, Victoria, South Australia, Northern Territory, Western Australia, or Tasmania), years working as a GP, mental health training (yes or no), and frequency of seeing adolescent patients (daily, weekly, fortnightly, monthly, or every 3-6 months). Participants were asked to rate the acceptability of the service (eg, how acceptable is this online clinic to you as a GP?) on a 5-point Likert scale ranging from unacceptable (1) to acceptable (5). Furthermore, 2 questions regarding frequency and type of feedback were asked (If you did refer a young person, how often would you like to receive feedback about your patient’s progress? and What feedback would you want to receive from the online clinic? [answered daily, weekly, fortnightly, monthly, or other]). GPs were asked what types of feedback notifications they would like to receive (registration, improvement, deterioration, suicidality, module completions, nonadherence, or other) and duration of follow-up (once your patient has recovered, how long would you like your patient to be followed up by the service? [answered 1-3 months, 6 months, or 12 months]) with the option of other (free response) to include alternate durations. Finally, GPs were asked if they had any other comments or advice (free response).

**Data Analysis**
The online survey was delivered using the Key Survey platform version 8.13, an online survey tool developed by WorldAPP. Data were then exported to IBM SPSS version 22, and basic descriptives were calculated and reported. Audio recordings of the interviews were transcribed and analyzed using the six phases of Braun and Clarke’s [34] thematic analysis guidelines. First, two researchers (MSK and MA) read through each transcript several times to gain an overview of the interview data before code generation. Second, all data extracts were initially coded, and the search for underlying themes began by combining the emerging codes. Third, all data extracts were classified in relation to the identified themes. This process was repeated for each transcript, and any clear patterns within the data were identified. The themes were then reviewed, and any data extracts that did not fit into the initial themes were once again investigated for identification of different themes. There was discussion on higher-order codes and on points of agreement or disagreement, leading to consensual validation. Final themes were agreed upon through collaborative analysis. Final coder agreement was 80%.

**Results**

**Participants**
A total of 13 GPs completed the telephone interview. Of these, 11 completed the online survey and 9 out of the 11 (82%) were female. Participants had a mean age of 51.0 years (standard deviation [SD] 12.16; range: 35-79). The mean number of years that participants had been working was 21 (SD 13.76; range: 4-50). All participants had undertaken additional mental health training. Nearly half of the participants (5/11, 46%) reported having weekly contact with adolescent patients, with less than one-third (3/11, 28%) reporting daily contact.

**Level of Acceptability of the Proposed Service**
All participants reported that the proposed service was entirely (9/11) or slightly (2/11) acceptable. Thematic analyses found that four key themes contributed to GPs’ acceptability of the proposed service (Table 1).
Tables 2-4 outline GPs’ preferences about the type and frequency of patient feedback and duration of follow-up within the proposed Web-based service.

As outlined in Table 5, eight themes emerged as potential barriers to the service. These themes were then classified into one of the four main issues: (1) barriers relating to the service model, (2) characteristics of youth patients, (3) environmental factors, and (4) GP characteristics.

Table 1. Themes contributing to the acceptability of the proposed service model (N=13).

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>n (%)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early intervention and prevention</td>
<td>The belief that the service provides an opportunity to detect youth with subthreshold symptoms who may remain undetected or receive little attention by traditional services, thus providing early intervention</td>
<td>7 (54)</td>
<td><em>I think it’s useful to be universally screening adolescents and it’s useful to be capturing this subsyndromal group early who generally would not receive any attention. I think the potential to intervene at that point and prevent further escalation is really important.</em> [GP1]</td>
</tr>
<tr>
<td>Adolescent preferences</td>
<td>The belief that the service was an appropriate platform for this age group because of adolescent preferences for technology and spending time online</td>
<td>8 (62)</td>
<td>...adolescents these days are completely plugged into technology and I think it’s giving them access to assessment and treatment via a platform that they’re comfortable with. [GP5]</td>
</tr>
<tr>
<td>School context</td>
<td>The belief that the service would lead to greater access of care because it is free, online, delivered in the school environment, and complements traditional modes of therapy (eg, face-to-face)</td>
<td>8 (62)</td>
<td>My experiences there are that school counselors are very busy...So, it’s great for them to be able to actually say well look here’s something that you can do in the meantime. [GP4]</td>
</tr>
<tr>
<td>Anonymity</td>
<td>The belief that this type of service would provide anonymity, reducing stigma and ensuring privacy</td>
<td>5 (39)</td>
<td><em>I think that they prefer it has a level of anonymity, which is great, and it doesn’t necessarily involve parents or care providers having to know what’s going on in your life, so I think it’s a great idea for young people.</em> [GP5]</td>
</tr>
</tbody>
</table>

Table 2. Type of patient feedback desired by general practitioners (N=11).

<table>
<thead>
<tr>
<th>Type of feedback notifications</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient registration to online service</td>
<td>9 (82)</td>
</tr>
<tr>
<td>Clinical improvement</td>
<td>10 (91)</td>
</tr>
<tr>
<td>Clinical deterioration</td>
<td>11 (100)</td>
</tr>
<tr>
<td>Suicidality</td>
<td>11 (100)</td>
</tr>
<tr>
<td>Number and type of modules completed</td>
<td>6 (55)</td>
</tr>
<tr>
<td>Ceasing use</td>
<td>10 (91)</td>
</tr>
<tr>
<td>Other: information about referral pathways, who the general practitioner has permission to speak to (eg, caregivers and health professionals); child protection or domestic violence issues; and details of specific areas covered in online modules</td>
<td>3 (27)</td>
</tr>
</tbody>
</table>

Table 3. Frequency of patient feedback desired by general practitioners (N=11).

<table>
<thead>
<tr>
<th>Frequency of feedback notifications</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>2 (18)</td>
</tr>
<tr>
<td>Weekly</td>
<td>0</td>
</tr>
<tr>
<td>Fortnightly</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Monthly</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Other: frequency should decrease over time; should be dependent on client progress or step allocation or symptom severity</td>
<td>3 (27)</td>
</tr>
</tbody>
</table>

Table 4. Duration of follow-up desired by general practitioners (N=11).

<table>
<thead>
<tr>
<th>Duration of follow-up</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 3 months</td>
<td>2 (18)</td>
</tr>
<tr>
<td>6 months</td>
<td>4 (36)</td>
</tr>
<tr>
<td>12 months</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Other: depends on duration of involvement and time taken to recover; should be determined by the clinical presentation and patient</td>
<td>2 (18)</td>
</tr>
</tbody>
</table>
Table 5. Potential barriers to the proposed service model (N=13). GPs: general practitioners. CBT: cognitive behavioral therapy.

<table>
<thead>
<tr>
<th>Issues and theme</th>
<th>Definition</th>
<th>n (%)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence and knowledge of eHealth</td>
<td>The degree to which GPs thought that there was a shortage of evidence-based online care and the personal preference for face-to-face treatment</td>
<td>11 (85)</td>
<td>I think one of the issues with online services for young people is a lack of evidence...GP13</td>
</tr>
<tr>
<td>Privacy and confidentiality</td>
<td>The degree to which GPs were concerned with data storage, access, and privacy between GPs, school counselors, and caregivers (eg, data being mishandled or accessed without authorization)</td>
<td>6 (46)</td>
<td>I think the important thing in that context is to make sure that there’s an area where students can go and access the program, maybe somewhere that’s quiet or separate and which respects their confidentiality as I think certainly confidentiality and privacy is a huge issue. GP7</td>
</tr>
<tr>
<td>Effectiveness and accuracy of the proposed model</td>
<td>The degree to which GPs were concerned about youth not receiving appropriate care</td>
<td>5 (38)</td>
<td>I guess the only concern one could ever have is that someone who’s severely unwell doesn’t end up with their matching care. GP11</td>
</tr>
<tr>
<td>Characteristics of youth patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noncompliance from adolescents</td>
<td>A concern that there may be a risk that some young people do not engage with the program and motivation to complete the online modules may be low</td>
<td>8 (62)</td>
<td>I have been aware of online CBT courses that are available, and my uptake is really bad when I suggested it to patients...adolescents, young adults; even when they’ve been enthusiastic, they don’t generally follow through with it. GP11</td>
</tr>
<tr>
<td>Environmental factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access issues</td>
<td>Potential constraints regarding availability of these services (eg, internet and phone access, and rural location) and client characteristics (eg, the inability to access service because of learning difficulties, cultural barriers, low school attendance, or complex clinical presentations)</td>
<td>5 (38)</td>
<td>Kids that live remote...their lack of access and connectivity, no Wi-Fi and a lot of kids have got phones but they’re all the old phones...not smartphones. GP8</td>
</tr>
<tr>
<td>Lack of services</td>
<td>Limited availability of appropriate services in rural and remote areas (eg, limited options for referral, lack of qualified professionals, and adolescent specific services)</td>
<td>4 (31)</td>
<td>I work in a rural location, and in reality, I have access to very, very few services and the services that are available are usually targeted to adults...GP8</td>
</tr>
<tr>
<td>GP characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Differences amongst GPs</td>
<td>Perceived differences in GPs’ mental health knowledge and experience with adolescents</td>
<td>5 (38)</td>
<td>With increased numbers of bulk-billing clinics, you’re not going to get GPs with complete GP training or mental health training...GP11</td>
</tr>
<tr>
<td>Noncompliance by GPs</td>
<td>Risk that GPs will not use service or have no incentive to use the service (ie, without rebate or ability to bulk bill for time spent outside direct consultation)</td>
<td>3 (23)</td>
<td>The remuneration you get isn’t there...because we can only charge Medicare when we’ve got the patient with us...so, you know when I ring the school afterward or when reading reports, writing reports, stuff like that—there isn’t a financial incentive. GP8</td>
</tr>
</tbody>
</table>

Likelihood of Referral and Integration Into Clinical Practice

All GPs reported that they were highly (7/11) or somewhat (4/11) likely to refer youth patients to the proposed Web-based service. Thematic analyses found that two themes influenced likelihood of referral (Table 6). All GPs reported they would be highly (6/11) or somewhat (5/11) likely to integrate the proposed Web-based service into their clinical practice. Thematic analyses revealed that six key needs influenced the likelihood of GPs integrating the proposed service (Table 7).
Table 6. Themes influencing likelihood of referral of the proposed Web-based service (N=13). GPs: general practitioners.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Definition</th>
<th>n (%)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived need</td>
<td>Whether the GP identified a need for the service</td>
<td>11 (85)</td>
<td>We’re all buckling under the strain with the amount of work to do not just with adolescents but with mental health in general and anything that will help management and prevention is a fantastic tool. [GP12]</td>
</tr>
<tr>
<td>Beliefs</td>
<td>Whether the GP felt the service would be helpful, promote help-seeking, or support clinical practice</td>
<td>4 (31)</td>
<td>It sounds like the system allows regular monitoring and safeguards to the GP, then it certainly helps the GP to really manage the person. [GP9]</td>
</tr>
</tbody>
</table>

Table 7. General practitioner (GP) needs for service integration (N=13).

<table>
<thead>
<tr>
<th>Need</th>
<th>Definition</th>
<th>n (%)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative approach to management of patient</td>
<td>A belief that the service was likely to be integrated if a coordinated approach between adolescents, caregivers, school counselors, GPs, and other health professionals is adopted</td>
<td>10 (77)</td>
<td>I think that a shared care model is excellent with everybody being on the same page is vital. [GP10]</td>
</tr>
<tr>
<td>Duty of care and medico-legal implications</td>
<td>The need for delineation about who is part of the team and who is responsible for responding to alerts. Informed consent, transparent information about client progress, and a user-friendly feedback system were also discussed. GPs also outlined the need for legal advice (eg, notifying parents, consent, and medico-legal responsibility) and mandatory reporting guides (eg, self-harm, underage sex, substance use, and domestic violence)</td>
<td>8 (62)</td>
<td>I think GPs are going to have questions like, so what am I going to do if I get a message from you that says the patient is deteriorating or that says the patient has stopped using the program; is it my medico-legal responsibility to follow the patient up?...those kinds of questions, GPs are going to want to know the answers to. [GP2]</td>
</tr>
<tr>
<td>Encourage relationship with GP</td>
<td>A belief that the program should encourage adolescents to build a relationship with a GP, for example, include information about how to find a GP, the types of issues a GP can assist with, and help adolescents with appointments</td>
<td>10 (77)</td>
<td>...having a teenager build a relationship with a GP almost needs to be a goal within this, to work toward them having their own GP maybe separate to their parents and to understand how to use their relationship with the GP... [GP1]</td>
</tr>
<tr>
<td>Quality of service information provided to GPs</td>
<td>The need for concise information provision, for example, specific details about how the service works and links to relevant information as a reference for GPs</td>
<td>8 (62)</td>
<td>...some sort of clear and detailed summary that explains some of the skills and the strategies that people are learning will help me to understand what it is that I’m referring people to. [GP7]</td>
</tr>
<tr>
<td>Community awareness and promotion of the service</td>
<td>Strategies to overcome a lack of community awareness, for example, promotion in schools, advertising in newspapers, and promotion in health centers by health professionals</td>
<td>11 (85)</td>
<td>...I’d need like some sort of way of referring the patients, so whether it’s sort of a brochure to hand over, or whether it’s card with the email or online link or whether it’s a referral pad. [GP10]</td>
</tr>
<tr>
<td>Training</td>
<td>A request for training for GPs and school counselors on how to use the service and provide feedback</td>
<td>6 (46)</td>
<td>I think an educational video would be really good... [GP7]</td>
</tr>
</tbody>
</table>

Discussion

Principal Findings

This study aimed to examine GPs’ attitudes toward a Web-based mental health service for treating depression and anxiety among adolescents. This study also aimed to identify barriers to integrating a Web-based service into clinical practice. This knowledge is important for determining best practice in the field of e-mental health service design and delivery for youth. In this study, most GPs reported that the proposed service was an acceptable type of care because of its focus on early intervention and prevention, the alignment with adolescents’ help-seeking preferences, the provision of anonymity to reduce stigma and ensure privacy, and its delivery in the school setting. These attitudes are consistent with other health workers, who have reported that access to care via the internet overcomes geographical, psychological, and physical barriers, is cost-effective, and a viable adjunct to standard therapy [1,15,16,35,36]. The sense of anonymity of online care somewhat overcomes issues relating to stigma and embarrassment about seeking help for mental health problems, which has previously been identified as the most prominent barrier to help-seeking in young people [37]. The stepped care
component of the model was also supported. GPs felt that the proposed service could improve the detection of mental health problems among youth and identify those with subthreshold symptoms who are currently not receiving care. Future evaluations of the service will need to provide evidence to support the effectiveness of such a model in this severity group.

Factors Influencing Likely Use

In this study, the GPs identified several issues that may influence the acceptability, both positively and negatively, of the proposed Web-based service. Participants reported that GPs may be concerned that young people using a Web-based service may not receive adequate care, or that some youth may not be able to access the service because of poor internet connectivity. In addition, participants reported that GPs may be worried about data security and privacy. Although this is a common concern of internet-delivered care [24,25,32], participants felt that this could be alleviated by clear delineation of roles and responsibilities within the service alongside robust informed consent procedures with the young person. Additional barriers to acceptability of the proposed service included noncompliance by adolescents and GPs, as well as differences in mental health training and youth experience of GPs. Interestingly, lack of financial incentives for extra consultation time was only mentioned by 3 GPs. This is in contrast to previous research that identified poor financial reimbursement as a major constraint to GPs delivering high-quality youth mental health care [10,13]. These differences in study findings may be because of unique characteristics of the Web-based service (eg, the school counselor being the primary carer, thus reducing the workload of the GP) or because of characteristics of the GPs interviewed (eg, all reported having mental health training, thus the service proposed may not be viewed as an additional burden on time). However, this is purely speculative and future studies would benefit from investigating whether financial incentives increase rates of referral among GPs.

Importance of Patient Follow-Up and Feedback

All GPs agreed that they would like feedback on symptom deterioration and suicidality, as well as clinical improvement. This is not surprising, given the significant ethical and legal obligations associated with providing treatment to adolescents. Interestingly, perspectives on the frequency of feedback and duration of follow-up varied. Half of the GPs surveyed wanted fortnightly or monthly feedback, whereas others felt feedback should be dependent on patient progress and/or symptom severity. Over half of the sample felt the duration of follow-up should not exceed 6 to 12 months. These results highlight the importance of monitoring functionality, and a future model of the service should incorporate customizable options for frequency of patient feedback and follow-up to account for GP preferences.

Likelihood of Referral and Integration Into Clinical Practice

All GPs reported that they were highly or somewhat likely to refer youth patients to the proposed Web-based mental health service. However, referral was found to be influenced by perceived need and beliefs, such that GPs who do not self-identify as needing assistance in providing care to youth, or those who do not believe in the effectiveness of Web-services, would be unlikely to refer to the proposed service. This is consistent with previous research in which GPs’ attitudes to treatments were largely influenced by their personal beliefs about effectiveness [26,33]. This poses a significant challenge to researchers and policy makers who are attempting to increase GPs’ use of e-mental health. A future trial of the proposed service may need to involve additional pretraining that addresses GP’s knowledge and awareness of e-mental health.

Although the participants reported that they would be likely to integrate the proposed service model into their clinical practice, six key needs were outlined. GPs were more likely to integrate the proposed service if it involved close collaboration with other health care professionals, duty of care and medico-legal implications were clearly addressed, the service encouraged an ongoing relationship with a GP, the quality of information provided to GPs about the service was concise, and if community awareness and promotion of the service, as well as training on how to use the service, was offered. Future trials of the service will need to ensure that GPs have mechanisms of maintaining contact with school counselors, parents, and any additional care providers involved in the treatment of the client. Trial studies of the service may benefit from demonstrations of the online service, alongside case studies and evaluation data to increase the likelihood of GPs recommending and using the service.

Limitations

Although this study provides support for the acceptability of a Web-based mental health service for adolescents by GPs across Australia, there were a few limitations. First, it is a small study in which only 13 participants were involved. These findings therefore present only an initial indication of GPs’ attitudes, which would be strengthened by a larger sample size. GPs in this study were not using the proposed service. Thus, the results indicate intention rather than actual behavior. In addition, all participants had undergone additional mental health training and had experience working with young people. Different results may have been found among GPs without this, and future work may be strengthened by targeting different GPs with less training and youth experience. Views may also differ among international samples, where the role of GPs in adolescent mental health care may vary.

Conclusions and Future Work

Overall, GPs in this study were open to the proposed service, but have concerns about certain aspects of Web-based care, characteristics of youth and GPs, and accessibility of a Web-based service. These would need to be addressed before GPs will refer or integrate the service into their clinical practice. In future work, researchers and service designers need to consider GPs as end users and evaluate the effects of the proposed service model on GPs’ confidence in delivering youth mental health care and the clinical effectiveness. Key next steps would be a pilot study of the service in primary care settings, alongside a formal evaluation of the effectiveness of the service for improving quality of care and symptom reductions among youth. Given the complex nature of the intervention and its
setting, the multiphase optimization strategy methodology [38], utilizing factorial designs, may be well suited to evaluating the independent effects of each of the components. This will assist the service designers and developers to optimize the service for both the practitioner and the patient.

Acknowledgments
Funding for this project was provided by Hong Kong and Shanghai Banking Corporation (HSBC).

Authors' Contributions
MSK was responsible for data analysis, interpretation, project management, and authorship of the manuscript. BOD and CK were involved in study design and manuscript preparation. MA assisted with data analysis. KOM undertook recruitment and data collection.

Conflicts of Interest
BOD is a Section Editor for *JMIR Mental Health*.

References
Abbreviations

- **CBT**: Cognitive Behavior Therapy
- **GP**: General practitioner
Subotic-Kerry M, King C, O’Moore K, Achilles M, O’Dea B
General Practitioners’ Attitudes Toward a Web-Based Mental Health Service for Adolescents: Implications for Service Design and Delivery
JMIR Hum Factors 2018;5(1):e12
URL: http://humanfactors.jmir.org/2018/1/e12/
doi:10.2196/humanfactors.8913
PMID:29572203

©Mirjana Subotic-Kerry, Catherine King, Kathleen O’Moore, Melinda Achilles, Bridianne O’Dea. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 23.03.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.
Three-Factor Structure of the eHealth Literacy Scale Among Magnetic Resonance Imaging and Computed Tomography Outpatients: A Confirmatory Factor Analysis

Lisa L. Hyde, BPsych (Hons); Allison W. Boyes, PhD; Tiffany-Jane Evans, MMed Stat; Lisa J. Mackenzie, PhD; Rob Sanson-Fisher, PhD

Health Behaviour Research Collaborative, School of Medicine and Public Health, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia
Priority Research Centre for Health Behaviour, University of Newcastle, Callaghan, Australia
Hunter Medical Research Institute, New Lambton Heights, Australia
Hunter Cancer Research Alliance, Newcastle, Australia

Corresponding Author:
Lisa L Hyde, BPsych (Hons)
Health Behaviour Research Collaborative, School of Medicine and Public Health
Faculty of Health and Medicine
University of Newcastle
Public Health/HBRC, HMRI Building, University of Newcastle
Callaghan, 2308
Australia
Phone: 61 249138799
Fax: 61 240420040
Email: Lisa.L.Hyde@uon.edu.au

Abstract

Background: Electronic health (eHealth) literacy is needed to effectively engage with Web-based health resources. The 8-item eHealth literacy scale (eHEALS) is a commonly used self-report measure of eHealth literacy. Accumulated evidence has suggested that the eHEALS is unidimensional. However, a recent study by Sudbury-Riley and colleagues suggested that a theoretically-informed three-factor model fit better than a one-factor model. The 3 factors identified were awareness (2 items), skills (3 items), and evaluate (3 items). It is important to determine whether these findings can be replicated in other populations.

Objective: The aim of this cross-sectional study was to verify the three-factor eHEALS structure among magnetic resonance imaging (MRI) and computed tomography (CT) medical imaging outpatients.

Methods: MRI and CT outpatients were recruited consecutively in the waiting room of one major public hospital. Participants self-completed a touchscreen computer survey, assessing their sociodemographic, scan, and internet use characteristics. The eHEALS was administered to internet users, and the three-factor structure was tested using structural equation modeling.

Results: Of 405 invited patients, 87.4% (354/405) were interested in participating in the study, and of these, 75.7% (268/354) were eligible. Of the eligible participants, 95.5% (256/268) completed all eHEALS items. Factor loadings were 0.80 to 0.94 and statistically significant (P<0.001). All reliability measures were acceptable (indicator reliability: awareness=.71-.89, skills=.78-.80, evaluate=.64-.79; composite reliability: awareness=.89, skills=.92, evaluate=.89; variance extracted estimates: awareness=.80, skills=.79, evaluate=.72). Two out of three goodness-of-fit indices were adequate (standardized root mean square residual (SRMR)=.038; comparative fit index (CFI)=.944; root mean square error of approximation (RMSEA)=.156). Item 3 was removed because of its significant correlation with item 2 (Lagrange multiplier [LM] estimate 104.02; P<.001) and high loading on 2 factors (LM estimate 91.11; P<.001). All 3 indices of the resulting 7-item model indicated goodness of fit ($\chi^2$=11.3; SRMR=.013; CFI=.999; RMSEA=.011).

Conclusions: The three-factor eHEALS structure was supported in this sample of MRI and CT medical imaging outpatients. Although further factorial validation studies are needed, these 3 scale factors may be used to identify individuals who could benefit from interventions to improve eHealth literacy awareness, skill, and evaluation competencies.
Introduction

Consumer eHealth Literacy is Critical to Maximizing the Benefits of eHealth

Technologically-enabled health care is important at both the patient and service level, given the increasing resource and timing pressures on the health care system [1], the digital transformation of health-related industries [2], and changing consumer expectations about their role in care [3]. Electronic health (eHealth) refers to the organization and delivery of health services and information using the internet and related technologies [4]. eHealth holds potential as a scalable form of service delivery that is accessible, low-cost, promotes patient empowerment, and enhances patient-provider information exchange [5]. However, to reap the possible benefits, patients must be eHealth literate [6]. eHealth literacy refers to an individual’s ability to seek, find, understand, and appraise health information from electronic sources, and apply the knowledge gained to addressing or solving a health problem [6]. Limited ability to seek, find, understand, and appraise electronic health information has been recognized as a key self-reported barrier to the utilization of the internet for health purposes [7]. The first step in identifying individuals who may benefit from improved eHealth literacy is the development of valid and reliable tools assessing this construct.

The eHealth Literacy Scale Is a Standardized and Widely Used Measure

The eHealth literacy scale (eHEALS) was among the first and continues to be one of the most commonly used self-reported measures of eHealth literacy [8,9]. The scale comprises 8 items, which assess consumers’ combined knowledge, comfort, and perceived skills at finding, evaluating, and applying electronic health information to health problems [8]. Consistent with the current definition of eHealth [4], all eHEALS items are specific to health information access via the Internet, as opposed to other electronic forms of information provision (eg, Compact Disc Read-Only Memory [CD-ROM], computer games). The scale was developed to address the need for an easily self-administrable eHealth literacy measure that could be applied across a wide range of populations and contexts [8]. Widespread adoption of the scale has been demonstrated, with the measure translated into multiple languages [10-17] and used across participants with diverse sociodemographic [10,15,16,18], ethnic [11,14,19], and disease profiles [13,20,21]. Items were originally developed and validated among Canadian youths more than a decade ago [8], and subsequent studies have demonstrated test-retest reliability across younger [14] and older age cohorts [10], internal consistency across populations of varying age and ethnicity [10,11,14,15,19,22], and measurement invariance across English-speaking countries [23]. However, inconsistent findings exist regarding the convergent and predictive validity of the scale [10-17,22,23,25-28]. We sought to contribute to this knowledge by assessing the factorial validity and internal consistency of a three-factor structure of the eHEALS.

The Factor Structure of the eHealth Literacy Scale Is Uncertain

Norman and Skinner’s original factorial validation of the eHEALS found that the scale assesses a single dimension [8]. Numerous studies with the general public have supported this finding [10,11,14-16,22,25,26], including those specific to children [15], university students [14,16], and older adults [10,22]. However, the strength of these conclusions is limited by the common use of exploratory factor analysis (EFA) [8,10,11,14,15,22,25,26]. EFA originates from classical test theory and holds value in the early stages of scale development when factor structure is unknown and latent variable structures need to be identified [29]. EFA does, however, have some limitations. For example, it often involves subjective decision-making processes and does not account for the theory which may inform latent variable structures [30].

Confirmatory factor analysis (CFA) is an alternative analysis technique, also derived from classical test theory, which allows models to be tested via theoretically or empirically-driven hypotheses [31]. However, studies assessing a unidimensional eHEALS structure using CFA commonly report poor fit indices [13,23,27,28]. This may be because a single factor structure does not account for the multifaceted nature of the concept of eHealth literacy, such as its inherent literacy types (ie, traditional, health, information, scientific, media, and computer) or the multiple components of information retrieval and use (ie, finding, applying and evaluating electronic health information) [6]. Paige and colleagues [13] completed one of the only studies of the construct validity of the eHEALS using CFA with chronically ill patients and found evidence for a three-factor structure. Despite this, multidimensionality of the eHEALS was refuted on the basis that a large proportion of variance loaded on one factor only. The authors applied the partial credit model, which is a unidimensional item response theory technique, to conclude that a single structure exists, despite CFA values indicating a poor unidimensional fit [13]. A two-factor model based on the concepts of information-seeking and appraisal has also been tested [12,27,28]. Although this model has a strong theoretical basis, 2 of the 3 studies testing this structure reported inadequate fit indices [12,27]. Furthermore, all were based on translated versions of the scale, which can result in varied item meaning and interpretation [32].

Recent Literature Proposes That the eHealth Literacy Scale Has a Three-Factor Structure

Sudbury-Riley and colleagues [23] used CFA to test a three-factor structure of the English-language version of the eHEALS with a multinational sample of adult internet users from the United Kingdom (n=407), New Zealand (n=276), and the United States (n=313). A hypothesis-driven approach was...
adopted, whereby 2 eHEALS items were mapped to an “awareness” factor, 3 items to a “skills” factor, and 3 items to an “evaluate” factor. These factors were derived from the self-efficacy and social-cognitive theoretical constructs underpinning eHealth literacy [8,23]. Self-efficacy theory is based on the premise that goal achievement is mediated by self-belief and confidence, and social cognitive theory states that social context influences goal achievement [33]. Sudbury-Riley and colleagues [23] therefore proposed that an individual’s awareness is shaped by their environment (eg, exposure to Web-based health information), their skills are influenced by social factors (eg, modeling, instruction, and social persuasion), and their ability to evaluate eHealth resources is mediated by their confidence and persistence. CFA fit indices supported the hypothesized three-factor eHEALS structure across all 3 countries [23].

Further Research Is Needed to Verify the Three-Factor Structure of the Standardized eHealth Literacy Scale With Patient Populations

Sudbury-Riley and colleagues’ [23] study contributes to our understanding of the underlying structures of the eHEALS, however, it has some limitations. In particular, a modified version of the scale was used, based on feedback from the authors’ family, friends, and colleagues, in which “and information” was added to items to address the increasing interactivity of eHealth materials. It is therefore unclear whether the three-factor structure also applies to the original version of the scale. The study was also conducted with middle-aged members of the general population, restricting the generalizability of findings across medical populations and age cohorts. This adds to the common underrepresentation of chronically ill patients in the eHEALS measurement literature, despite the potential benefits of eHealth to this population [13].

Given that evidence about the properties of a measure is accumulated over a number of studies, the appropriate next step it is to determine whether Sudbury-Riley and colleagues’ [23] findings can be replicated in a different population. To address this need, and also overcome some of the limitations of Sudbury-Riley and colleagues’ work [23], this factorial validation study was conducted with patients, using the standardized eHEALS. Magnetic resonance imaging (MRI) and computed tomography (CT) medical imaging outpatients represent a high volume of patients with diverse demographic characteristics and medical diagnoses [34,35], and as such, research completed with these patients may have high generalizability. Furthermore, MRI and CT medical imaging outpatients require substantial preparatory information that could potentially be delivered online [36]. Hence, this study aimed to test the factorial validity and internal consistency of the three-factor structure of the eHEALS, identified by Sudbury-Riley and colleagues [23], among MRI and CT medical imaging outpatients.

Methods

Design and Setting

A cross-sectional survey of CT and MRI medical imaging outpatients was conducted in a medical imaging clinic at a tertiary referral hospital located in regional New South Wales, Australia.

Participants

Eligible participants were attending for an outpatient CT or MRI appointment at the tertiary referral hospital, were 18 years or older, and had access to the internet for personal use. Participants were excluded from the study if they had a cognitive or physical impairment that precluded them from providing informed consent or participating in the study, or if they were unable to complete the questionnaire because of poor English proficiency. These criteria mean that a diversity of participants in terms of frequency, confidence, and reasons for personal use of the internet were eligible to participate. Consistent with the original eHEALS validation study [8], use of the internet for health was not an eligibility requirement.

Procedure

Patients who were potentially eligible for the study were identified by medical imaging reception staff when they presented for their outpatient appointment. These patients were informed about the research and invited to speak with a trained research assistant. Interested patients were provided with a written information sheet and introduced to the research assistant, who gave an overview of the study and obtained patients’ verbal consent to participate. The age, gender, and scan type of noninterested and nonconsenting patients were recorded. Consenting participants were provided with a tablet computer and asked to complete a Web-based questionnaire before their scan. A paper version of the questionnaire was provided to participants who requested it. Ethics approval was obtained from the Hunter New England Human Research Ethics Committee (16/10/19/5.11) and University of Newcastle (H-2016-0386).

Measures

Participants’ eHealth literacy was assessed using the 8-item English-language version of the eHEALS [8]. Respondents indicated their level of agreement with each statement on a 5-point Likert scale, which was scored from 1 “strongly disagree” to 5 “strongly agree.”

Sociodemographic, scan, and information preference characteristics were examined using standard items. These items assessed participant age, gender, marital status, highest level of education completed, postcode, and scan type. Postcodes were mapped to the Accessibility/Remoteness Index of Australia Plus 2011 classification to examine remoteness [37] and categorized as metropolitan (major cities of Australia) or nonmetropolitan (inner regional, outer regional, remote, or very remote Australia). One item, adapted from an existing health information wants questionnaire [38], assessed how much information participants liked to have about their health.
Response options were “no information,” “some information,” and “a lot of information.”

**Figure 1.** eHealth Literacy Scale three-factor model proposed by Sudbury-Riley and colleagues.

![eHealth Literacy Scale three-factor model](image)

Internet characteristics were assessed by 2 items. Use of the internet for scan preparation was assessed by an author-developed item: *Have you searched the internet for information to help you prepare for your scan?* with response options “no,” “yes,” and “don’t know.” Frequency of internet use was measured with a single item used in existing informatics literature [39], in which participants respond on a 6-point scale ranging from “less than once a month” to “several times a day.”

**Sample Size**

Rules of thumb for CFA recommend a sample size of at least 200 participants [40,41] or 10 participants per parameter estimated [42]. Wolf and colleagues [43] found that a sample size of at least 150 is required for three-factor models with fewer than 4 indicator variables per factor and assuming strong factor loadings of 0.80. To accommodate deviation from these assumptions, and given that 19 parameters were estimated for the eHEALS CFA, the more conservative estimate of at least 200 participants was applied to this study.

**Statistical Analyses**

Participant characteristics and eHEALS responses were summarized as frequencies and percentages, or means and standard deviations. Consent bias was assessed for gender, scan type, and age group using chi-square tests. CFA was undertaken using the CALIS procedure of SAS software v9.4 (SAS Institute, Cary, NC, USA). We chose CFA as it is the same theoretically-sound technique used by Sudbury-Riley and colleagues [23] and therefore allowed for a direct comparison of results. Given the high completion rate (98.1% [256/261] of participants who started the eHEALS completed all items), this analysis was restricted to participants with complete eHEALS data. The relationship between latent variables (ie, awareness, skills, evaluate) and manifest variables (eHEALS items 1-8), as proposed by Sudbury-Riley and colleagues [23], was tested using structural equation modeling (Figure 1). All loadings were standardized, with variances fixed at 1. The model was estimated using the full information maximum likelihood method. Standardized factor loadings and covariances were calculated with 95% CIs.

Reliability measures included indicator reliability to determine the percentage of variation in the item explained by each factor, composite reliability to assess internal consistency (> .70 ideal) [29], and variance extracted estimates (VEEs) to determine the amount of variance captured by factors with regard to variance attributable to measurement error (> .50 ideal) [44]. Discriminant validity was assessed following the method proposed by Anderson and Girbing [45].

Model goodness of fit was assessed using a range of metrics. Absolute indices included the chi-square statistic, the chi-square to degrees of freedom ratio (< 2 ideal) [46], and the standardized root mean square residual (SRMR; < .05 ideal) [29]. The incremental index was reported as the comparative fit index (CFI; > .95 good fit) [47]. The parsimony index used was the root mean square error of approximation (RMSEA; < .05 close approximation fit, .05-.08 acceptable fit, > .10 poor fit) [29,47]. Lagrange multiplier (LM) estimates of items on different factors were assessed to identify complex items and possible ways to improve the model.

**Results**

**Sample**

A total of 405 potentially eligible patients were invited to discuss the study with a research assistant during the 7-week recruitment period. Of the invited patients, 87.4% (354/405) were interested in participating in the study, and of these, 75.7% (268/354) were eligible. Of these eligible participants, 97.4% (261/268) started
the eHEALS, and 95.5% (256/268) completed all eHEALS items. There were no significant differences between patients who were and were not interested in participating in the study based on gender, scan type, or age group. Table 1 provides a summary of the sociodemographic, scan, and internet characteristics of eligible participants. Multimedia Appendix 1 provides a summary of participant responses to eHEALS items.

**Confirmatory Factor Analysis**

Convergence between the implied and observed variance covariance matrices was achieved within 10 iterations. As shown in Table 2, all factor loadings were at or above 0.80 and were statistically significant ($P < .001$). All CRs exceeded .70, indicating good reliability, and all VEEs exceeded the cutoff of .50, indicating convergent validity. Discriminant validity of the model was demonstrated, with statistically significant chi-square difference-tests ($P < .001$) for each pair of factors. The absolute index SRMR was .038, indicating adequate fit to the hypothesized model. The incremental index CFI was .944 and therefore close to the .95 threshold of acceptability (Table 3). However, the chi-square statistic ($\chi^2_{17} = 124.2$) was highly significant and suggestive of poor fit, and the chi-square statistic to degrees of freedom ratio of 7.3 exceeded the acceptability cutoff of 2 [46]. The parsimony index RMSEA was .16, indicating poor fit.
Table 1. Participant sociodemographic, scan, and internet characteristics (N=268).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n (%)^a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age years (SD)</td>
<td>53 (15)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>120 (44.8)</td>
</tr>
<tr>
<td>Female</td>
<td>148 (55.2)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Married or partner</td>
<td>148 (64.9)</td>
</tr>
<tr>
<td>Not married/living with partner</td>
<td>80 (35.1)</td>
</tr>
<tr>
<td>Education completed</td>
<td></td>
</tr>
<tr>
<td>High school or less</td>
<td>169 (63.1)</td>
</tr>
<tr>
<td>More than high school</td>
<td>99 (36.9)</td>
</tr>
<tr>
<td>Geographic location</td>
<td></td>
</tr>
<tr>
<td>Metropolitan</td>
<td>212 (79.1)</td>
</tr>
<tr>
<td>Nonmetropolitan</td>
<td>56 (20.9)</td>
</tr>
<tr>
<td>Scan type</td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>104 (38.8)</td>
</tr>
<tr>
<td>MRI</td>
<td>160 (59.7)</td>
</tr>
<tr>
<td>Don’t know</td>
<td>4 (1.5)</td>
</tr>
<tr>
<td>Used internet for scan</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>29 (10.9)</td>
</tr>
<tr>
<td>No</td>
<td>237 (88.8)</td>
</tr>
<tr>
<td>Don’t know</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td>Frequency of internet use</td>
<td></td>
</tr>
<tr>
<td>Less than once a month</td>
<td>11 (4.1)</td>
</tr>
<tr>
<td>Once a month</td>
<td>5 (1.8)</td>
</tr>
<tr>
<td>A few times a month</td>
<td>14 (5.2)</td>
</tr>
<tr>
<td>A few times a week</td>
<td>36 (13.5)</td>
</tr>
<tr>
<td>About once a day</td>
<td>51 (19.1)</td>
</tr>
<tr>
<td>Several times a day</td>
<td>150 (56.2)</td>
</tr>
<tr>
<td>Information amount preference</td>
<td></td>
</tr>
<tr>
<td>No information</td>
<td>2 (0.8)</td>
</tr>
<tr>
<td>Some information</td>
<td>59 (26.0)</td>
</tr>
<tr>
<td>A lot of information</td>
<td>166 (73.1)</td>
</tr>
</tbody>
</table>

^aNumber of observations for each characteristic may not total 268 because of missing data.
Table 2. Factor loading and residual error estimates for confirmatory factor analysis of hypothesized model.

<table>
<thead>
<tr>
<th>Factor-variable</th>
<th>Factor loadings (95% CI)</th>
<th>Error estimates (95% CI)</th>
<th>IR</th>
<th>CR</th>
<th>VEE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Awareness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I know what health resources are available on the Internet</td>
<td>0.85 (0.80-0.89)</td>
<td>0.29 (0.21-0.36)</td>
<td>0.71</td>
<td>0.92</td>
<td>0.79</td>
</tr>
<tr>
<td>I know where to find helpful health resources on the Internet</td>
<td>0.94 (0.91-0.97)</td>
<td>0.11 (0.05-0.17)</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td><strong>Skills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I know how to find helpful health resources on the Internet</td>
<td>0.90 (0.86-0.93)</td>
<td>0.20 (0.14-0.26)</td>
<td>0.80</td>
<td>0.92</td>
<td>0.79</td>
</tr>
<tr>
<td>I know how to use the internet to answer my questions about health</td>
<td>0.88 (0.85-0.92)</td>
<td>0.22 (0.16-0.28)</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>I know how to use the information I find on the internet to help me</td>
<td>0.88 (0.85-0.92)</td>
<td>0.22 (0.16-0.28)</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Evaluate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have the skill I need to evaluate the health resources I find on the Internet</td>
<td>0.85 (0.80-0.89)</td>
<td>0.21 (0.15-0.28)</td>
<td>0.79</td>
<td>0.80</td>
<td>0.72</td>
</tr>
<tr>
<td>I can tell high quality from low quality health resources on the Internet</td>
<td>0.86 (0.82-0.90)</td>
<td>0.26 (0.19-0.33)</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>I feel confident in using information from the internet to make health decisions</td>
<td>0.80 (0.75-0.85)</td>
<td>0.36 (0.28-0.44)</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
</tr>
</tbody>
</table>

aIR: indicator reliability.
bCR: composite reliability.
cVEE: variance extracted estimate.
dP<.001.
eThis item was dropped in the alternative 7-item model.

Table 3. Goodness-of-fit indices for tested models.

<table>
<thead>
<tr>
<th>Index type and fit index</th>
<th>Statistics for hypothesized 8-item model</th>
<th>Statistics for tested 7-item model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-square</td>
<td>124.2</td>
<td>11.3</td>
</tr>
<tr>
<td>Chi-square degrees of freedom</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>P-value for the chi-square statistic</td>
<td>&lt;.001</td>
<td>.417</td>
</tr>
<tr>
<td>SRMR</td>
<td>.038</td>
<td>.012</td>
</tr>
<tr>
<td>Incremental index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bentler CFI</td>
<td>.944</td>
<td>.999</td>
</tr>
<tr>
<td>Parsimony index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEA estimate</td>
<td>.156</td>
<td>.011</td>
</tr>
<tr>
<td>RMSEA upper 90% CI</td>
<td>.131</td>
<td>.000</td>
</tr>
<tr>
<td>RMSEA lower 90% CI</td>
<td>.182</td>
<td>.066</td>
</tr>
</tbody>
</table>

aSRMR: standardized root mean square residual.
bCFI: comparative fit index.
cRMSEA: root mean square error of approximation.

When investigating the possible reasons for less than ideal fit, LM estimates provided strong evidence for a path between item 3 “I know how to find helpful health resources on the Internet” and the awareness factor (LM estimate 107.66; P<.001). There was also strong evidence for a path between item 2 “I know where to find helpful health resources on the Internet” and item 3 “I know how to find helpful health resources on the Internet” (LM estimate 91.11; P<.001). Given apparent overlap between items 2 and 3, a 7-item model which excluded item 3 was tested, which indicated good model fit (Table 3). See Multimedia Appendix 2 for factor loading and residual error estimates for this altered model.

Discussion

Principal Findings

This study was the first to examine the theoretically-derived three-factor structure of the eHEALS, as proposed by Sudbury-Riley and colleagues [23], among a sample of MRI

and CT medical imaging outpatients. This three-factor structure was supported, with 2 out of 3 goodness-of-fit indices indicating adequate fit to the hypothesized model. Although these findings oppose accumulated evidence for a unidimensional structure of the eHEALS [8,10,11,14-16,22,25,26], they are consistent with the social cognitive and self-efficacy theory underpinning eHealth literacy [8,23,33]. As a result, it may be timely for researchers to examine patients’ eHealth literacy across eHEALS factors to inform targeted eHealth literacy improvement interventions. This study contributes important knowledge about the structure of the eHEALS, yet further factorial analyses, including multidimensional item response theory analyses, are required across populations to increase the reliability of these findings.

**Findings Broadly Support the Proposed Three-Factor Structure of the eHEALS**

The proposed model demonstrated strong internal consistency and discriminant validity, suggesting that items within each factor measured the same general construct, and these constructs were sufficiently different from one another. Similarly, 2 out of 3 fit indices demonstrated good fit to the proposed three-factor model. Factor loadings were high and statistically significant, similar to that reported by Sudbury-Riley and colleagues [23]. This finding contrasts to the majority of existing literature, where it is argued that a single factor structure exists [8,10-16,19,22,25,26]. Most such prior research is based on data-driven EFA techniques [8,10,11,14,15,22,25,26], which may indicate that limited reference to the theoretical underpinnings of eHealth literacy has resulted in inaccurate interpretations of eHEALS data in the past.

**Not all Goodness-of-Fit Indices Were Ideal**

Poor fit of the parsimony index suggests that complexity exists within the three-factor model. RMSEA estimates have also been identified as a poor performing goodness-of-fit metric in other CFA eHEALS literature [12,13,27] and are rarely reported as being a close approximate fit, indicating that relationships among items need to be interrogated. When we investigated further, it was found that item 3 “I know how to find helpful health resources on the Internet” loaded on both “skills” and “awareness” domains, and correlated significantly with item 2 “I know where to find helpful health resources on the Internet.”

This finding supports that of Sudbury-Riley and colleagues [23], who identified substantial overlap between items 2 and 3. Potential item homogeneity is also evident in prior literature, as measures of internal consistency have commonly been reported to be approaching the .95 threshold of acceptability for Cronbach alpha [10,11,15,19], with some reported to have reached .97 [22]. The redundancy of items 2 and 3 is unsurprising, given their similar structure and meaning (ie, about how and where to find helpful health resources on the Internet). It is also possible that the low education level of the sample [48], and the distressing setting of a hospital waiting room [49], contributed to participants’ difficulties in differentiating between item meanings. However, patient understanding of eHEALS items has been questioned previously, and the need for further research investigating item interpretation across populations has been indicated [11].

For this study, we did not restrict our sample to health-related internet users. This aligns with the majority of studies assessing the factorial validity of the eHEALS, including Norman and Skinner’s original validation study [8,10-17,19,22,26-28]. Furthermore, Norman and Skinner [8] highlight the potential application of the scale to those with varying levels of technology use. eHEALS response options of disagree and strongly disagree provide for those who do not use the internet for health. Despite this, some participants within this study voluntarily reporting being unsure of how to respond to each item as they did not use the internet for health. This anecdotal feedback suggests that items within the scale may not be interpretable to the wide population for which it was originally intended [8], and further research is needed to investigate the face and content validity of the scale among those who do and do not use the internet for health purposes.

As model fit improved when item 3 “I know how to find helpful health resources on the Internet” was excluded, an adapted 7-item eHEALS may be appropriate to consider. Reducing the number of items would result in two factors containing 2 items, which could create difficulties with model identification and convergence [29]. Likewise, it is unknown whether a reduced 2-item “skill” factor would adequately measure the construct and appropriately detect changes over time. As such, further research is needed to test the psychometric properties (specifically content validity, test-retest reliability, predictive validity, and responsiveness) of a 7-item eHEALS. Until this point, it is recommended that the standardized 8-item scale is used, with consideration of preliminary evidence supporting a three-factor structure.

**The Three-Factor Structure of the eHEALS May Reflect an eHealth Literacy Pathway Among internet Users**

Despite some fit indices being less than ideal, considering eHealth literacy by factor may help to guide Web-based health information provision in research and clinical practice. Furthermore, in accordance with the eHealth literacy continuum proposed by Diviana and colleagues [12], the eHEALS may measure an eHealth literacy pathway. In this instance, eHEALS factors are structured sequentially, and a user gradually demonstrates proficiency in more complex tasks. That is, a user must first be aware of eHealth resources before they can use their skills to navigate and interact with electronic content, and finally evaluate content quality and applicability to their health situation. Only once a user has undertaken all 3 of these steps, will they be able to effectively engage with eHealth resources and reap related benefits. This proposed pathway structure is supported by findings of Neter and colleagues [24], who reported that success rates gradually declined for older adults performing health-related computerized simulation tasks, as they stepped through the process of accessing, understanding, appraising, applying, and generating new health information. These findings may, however, be influenced by order effects of the simulated tasks [50], and further research is needed to validate such a causal pathway.
Important Implications for the Future Development and Evaluation of eHealth Literacy Improvement Strategies

On the basis of these findings, researchers and health care professionals have the opportunity to identify areas (ie, awareness, skills, or evaluate) where competency is low and target eHealth literacy improvement interventions accordingly. These interventions may, for example, include clinician recommendations to Web-based materials to increase awareness and reduce the need to evaluate content [51], training sessions to enhance eHealth literacy skills [52], or the promotion of checklists to aid in the evaluation of Web-based resources [53]. Additionally, user characteristics such as sociodemographic, health, and Internet use attributes that are associated with lower competency across eHEALS factors could be identified, so that assistance is directed toward those most in need. No studies have been conducted to determine the competency of individuals across eHEALS awareness, skill, and evaluate domains, and further research is needed.

Limitations

CFA was selected as it represents an understudied yet rigorous aspect of classical test theory and logically extends on the existing body of EFA and CFA measurement literature. The recent emergence of item response theory analyses of the eHEALS [12,13,16] has advantages over classical test theory approaches, including the capacity to establish increased item level psychometric information (eg, item difficulty). The application of multidimensional item response theory techniques to validate the three-factor eHEALS structure should be explored further. Furthermore, this study assessed one psychometric property (ie, factorial validity), and more research is needed to investigate other understudied measurement properties of the eHEALS, such as its predictive validity.

It is possible that findings may not be generalizable beyond the medical imaging context. Similarly, as most participants reported using the internet at least daily (75.3%, 201/267), study findings may not be generalizable to those who use the internet less frequently. As we did not ask participants about the activities they undertook online, it is unclear whether the results are applicable to those who do or do not use the internet for health. Future research is consequently needed to validate study findings across patients with diverse demographics, medical diagnoses, and internet use patterns. Additionally, our study was based on the standardized version of the eHEALS. As recognized in prior research [12,23], this version may not sufficiently capture competency in using Web 2.0 (eg, social networking) for health. Further research is needed to determine whether scale modifications are needed to reflect the evolving nature of eHealth interventions.

Conclusions

Although potential item redundancy impacted fit indices, the three-factor structure of the eHEALS was broadly supported. On the basis of these findings, the eHEALS could be used to inform the development of tailored eHealth literacy enhancement strategies, which may in turn increase engagement with Web-based health resources. Further research is needed to confirm the three-factor structure across other medical settings and populations to support the generalizability of these findings.

Acknowledgments

The authors thank the patients for their involvement in this study, as well as the administrative and clinical staff at Hunter New England Imaging for assistance with recruitment and data collection. This research was supported by a Hunter Cancer Research Alliance statistical support grant (HCRA 157) and has received funding support from the Hunter Cancer Research Alliance Implementation Science Flagship Program as part of the 2018 HDR Student Award initiative. LH is supported by an Australian Government Research Training Program Scholarship. AB is supported by a National Health and Medical Research Council Early Career Fellowship (APP1073317) and Cancer Institute New South Wales Early Career Fellowship (13/ECF/1-37). LM is supported by a Postdoctoral Fellowship (PF-16-011) from the Australian National Breast Cancer Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Participant responses to eHEALS items (n=261).

[PDF File (Adobe PDF File), 26KB - humanfactors_v5i1e6_app1.pdf]

Multimedia Appendix 2

Factor loading and residual error estimates for the confirmatory factor analysis of the 7-item model.

[PDF File (Adobe PDF File), 31KB - humanfactors_v5i1e6_app2.pdf]

References


Abbreviations

CFA: confirmatory factor analysis  
CFI: comparative fit index  
CT: computed tomography  
EFA: exploratory factor analysis  
LM: Lagrange multiplier  
eHEALS: 8-item eHealth literacy scale  
eHealth: electronic health  
MRI: magnetic resonance imaging  
RMSEA: root mean square error of approximation  
SRMR: standardized root mean square residual  
VEE: variance extracted estimate