JMIR Human Factors

Impact Factor (2022): 2.7
Volume 8 (2021), Issue 4 ISSN 2292-9495 Editor in Chief: Andre Kushniruk, BA, MSc, PhD, FACMI

Contents

Original Papers

User-Centered System Design for Communicating Clinical Laboratory Test Results: Design and Evaluation Study (e26017)
Zhan Zhang, Lukas Kmoth, Xiao Luo, Zhe He. . . . 4

Evaluating An Automated Compounding Workflow Software for Safety and Efficiency: Implementation Study (e29180)
Ülle Meren, James Waterson. . . . 19

Using Postmarket Surveillance to Assess Safety-Related Events in a Digital Rehabilitation App (Kaia App): Observational Study (e25453)
Deeptee Jain, Kevin Norman, Zachary Werner, Bar Makovoz, Turner Baker, Stephan Huber. . . . 30

A Novel Digital Pill System for Medication Adherence Measurement and Reporting: Usability Validation Study (e30786)
Susan Baumgartner, D Buffkin Jr, Elise Rukavina, Jason Jones, Elizabeth Weiler, Tony Carnes. . . . 48

The Effectiveness of a Multidisciplinary Electronic Discharge Readiness Tool: Prospective, Single-Center, Pre-Post Study (e27568)
Angela Keniston, Lauren McBeth, Jonathan Pell, Kasey Bowden, Anna Metzger, Jamie Nordhagen, Amanda Anthony, John Rice, Marisha Burden. . . . 60

Computer-Aided Screening of Autism Spectrum Disorder: Eye-Tracking Study Using Data Visualization and Deep Learning (e27706)
Federica Cilia, Romuald Carette, Mahmoud Elbattah, Gilles Dequen, Jean-Luc Guérin, Jérôme Bosche, Luc Vandromme, Barbara Le Driant. 7 7

Perceptions of Patients and Physicians on Teleconsultation at Home for Diabetes Mellitus: Survey Study (e27873)
Nazaré Rego, Helena Pereira, José Crispim. . . . 88

User Interactions With Health Insurance Decision Aids: User Study With Retrospective Think-Aloud Interviews (e27628)
Wayne Giang, Emma Bland, Jeffrey Chen, Coralys Colón-Morales, Michelle Alvarado. . . . 104

User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service (e30867)
Hedvig Aminoff, Sebastiaan Meijer, Kristina Groth, Urban Arnelo. . . . 116

Clinician Preimplementation Perspectives of a Decision-Support Tool for the Prediction of Cardiac Arrhythmia Based on Machine Learning: Near-Live Feasibility and Qualitative Study (e26964)
Stina Matthiesen, Søren Diederichsen, Mikkel Hansen, Christina Villumsen, Mats Lassen, Peter Jacobsen, Niels Risum, Bo Winkel, Berit Philbert, Jesper Svendsen, Tariq Andersen. 130

Pediatric Clinicians' Use of Telemedicine: Qualitative Interview Study (e29941)
Julia Finkelstein, Elise Tremblay, Melissa Van Cain, Aaron Farber-Chen, Caitlin Schumann, Christina Brown, Ankoor Shah, Erinn Rhodes. . .
1 4 7

A Shared Decision-making Tool for Drug Interactions Between Warfarin and Nonsteroidal Anti-inflammatory Drugs: Design and Usability Study (e28618)
Thomas Reese, Guilherme Del Fiol, Keaton Morgan, Jason Hurwitz, Kensaku Kawamoto, Ainhoa Gomez-Lumbreras, Mary Brown, Henrik Thiess, Sara Vazquez, Scott Nelson, Richard Boyce, Daniel Malone. 158

Medical Specialists' Perspectives on the Influence of Electronic Medical Record Use on the Quality of Hospital Care: Semistructured Interview Study (e27671)
Rube van Poelgeest, Augustinus Schrijvers, Albert Boonstra, Kit Roes. 168

Stakeholder Perspectives on an Inpatient Hypoglycemia Informatics Alert: Mixed Methods Study (e31214)
Nestoras Mathioudakis, Moeen Aboabdo, Mohammed Abusamaan, Christina Yuan, LaPricia Lewis Boyer, Scott Pilla, Erica Johnson, Sanjay Desai, Amy Knight, Peter Greene, Sherita Golden. 178

Exploring Usability Issues of a Smartphone-Based Physician-to-Physician Teleconsultation App in an Orthopedic Clinic: Mixed Methods Study (e31130)
Songphan Choemprayong, Chris Charoenlap, Krerk Piromsopa. 193

An Embodied Conversational Agent in an eHealth Self-management Intervention for Chronic Obstructive Pulmonary Disease and Chronic Heart Failure: Exploratory Study in a Real-life Setting (e24110)
Silke ter Stal, Joanne Sloots, Aniel Ramlal, Harm op den Akker, Anke Lenferink, Monique Tabak. 209

A Mobile App to Support Self-management of Chronic Kidney Disease: Development Study (e29197)
Talar Markossian, Jason Boyda, Jennifer Taylor, Bella Etingen, François Modave, Ron Price, Holly Kramer. 225

Perceptions of Older Men Using a Mobile Health App to Monitor Lower Urinary Tract Symptoms and Tamsulosin Side Effects: Mixed Methods Study (e30767)
Elizabeth Wang, Benjamin Breyer, Austin Lee, Natalie Rios, Akinyemi Oni-Orisan, Michael Steinman, Ida Sim, Stacey Kenfield, Scott Bauer. .
2 3 6

Usability of an At-Home Anterior Nares SARS-CoV-2 RT-PCR Sample Collection Kit: Human Factors Feasibility Study (e29234)
Laura Strong, Irene Middendorf, Michelle Turner, David Edwards V, Varun Sama, Joshua Mou, K Adams. 246

Digitally Disconnected: Qualitative Study of Patient Perspectives on the Digital Divide and Potential Solutions (e33364)
Maria Alkureishi, Zi-Yi Choo, Ali Rahman, Kimberly Ho, Jonah Benning-Shorb, Gena Lenti, Itzel Velázquez Sánchez, Mengqi Zhu, Sachin Shah, Wei Lee. 257

Communicating Guideline Recommendations Using Graphic Narrative Versus Text-Based Broadcast Screensavers: Design and Implementation Study (e27171)
Lauren Sinnenberg, Craig Umscheid, Frances Shofer, Damien Leri, Zachary Meisel. 276

Review

Including the Reason for Use on Prescriptions Sent to Pharmacists: Scoping Review (e22325)
Kathryn Mercer, Caitlin Carter, Catherine Burns, Ryan Tennant, Lisa Guirguis, Kelly Grindrod. 41

Original Paper

User-Centered System Design for Communicating Clinical Laboratory Test Results: Design and Evaluation Study

Zhan Zhang[1*], PhD; Lukas Kmoth[1*], BA; Xiao Luo[2*], PhD; Zhe He[3*], PhD

[1]School of Computer Science and Information Systems, Pace University, New York, NY, United States

[2]School of Engineering & Technology, Indiana University–Purdue University Indianapolis, Indianapolis, IN, United States

[3]School of Information, Florida State University, Tallahasse, FL, United States

[*]all authors contributed equally

Corresponding Author:
Zhan Zhang, PhD
School of Computer Science and Information Systems
Pace University
163 William Street
New York, NY, 10078
United States
Phone: 1 2123461897
Email: zzhang@pace.edu

Abstract

Background: Personal clinical data, such as laboratory test results, are increasingly being made available to patients via patient portals. However, laboratory test results are presented in a way that is difficult for patients to interpret and use. Furthermore, the indications of laboratory test results may vary among patients with different characteristics and from different medical contexts. To date, little is known about how to design patient-centered technology to facilitate the interpretation of laboratory test results.

Objective: The aim of this study is to explore design considerations for supporting patient-centered communication and comprehension of laboratory test results, as well as discussions between patients and health care providers.

Methods: We conducted a user-centered, multicomponent design research consisting of user studies, an iterative prototype design, and pilot user evaluations, to explore design concepts and considerations that are useful for supporting patients in not only viewing but also interpreting and acting upon laboratory test results.

Results: The user study results informed the iterative design of a system prototype, which had several interactive features: using graphical representations and clear takeaway messages to convey the concerning nature of the results; enabling users to annotate laboratory test reports; clarifying medical jargon using nontechnical verbiage and allowing users to interact with the medical terms (eg, saving, favoriting, or sorting); and providing pertinent and reliable information to help patients comprehend test results within their medical context. The results of a pilot user evaluation with 8 patients showed that the new patient-facing system was perceived as useful in not only presenting laboratory test results to patients in a meaningful way but also facilitating in situ patient-provider interactions.

Conclusions: We draw on our findings to discuss design implications for supporting patient-centered communication of laboratory test results and how to make technology support informative, trustworthy, and empathetic.

(JMIR Hum Factors 2021;8(4):e26017) doi:10.2196/26017

KEYWORDS

clinical laboratory results; patient-centered care; patient portal; health communication

Introduction

Motivation

Health care organizations are increasing direct access of patients to their clinical data via patient portals [1-3]. For example, patients can check their laboratory test results, an important type of medical record data, on the portals outside of the clinical environment (eg, at home). It has been shown that providing patients with access to such data can lead to better patient-centered medical care [1,4-6] and enhanced patient-provider relationships [7,8]. Despite these potential benefits, the literature points out that merely providing access

to laboratory test results is insufficient to improve patient engagement in their care because many patients are not able to make sense of the data and, as such, their use of laboratory test results is significantly limited [9-11].

Prior work has argued that the core issue is not that patients do not have the ability to understand test results but that the current design of patient portals inhibits effective result comprehension [12]. That is, many portals only present test results to patients in a table, which was originally formatted for clinician interpretation [11,13,14]. As a result, lay individuals, especially those with low health literacy and numeracy, have difficulty identifying meaningful information from their laboratory test results, such as how concerning the results are and what they should do to cope with them [15]. It is therefore imperative to ensure that patients not only have access to their laboratory test results but are also able to *understand* and *act upon* them.

To achieve this goal, the key challenge to address when communicating laboratory test results to patients is how to optimize the way the results are presented [10] and provide patients with the necessary information to comprehend each result [12,16]. However, research on these aspects is limited. Only a few studies have explored the perceptions of patients on viewing laboratory test results via portals [12,17,18] and what visual cues might be useful for aiding result comprehension [19-22]. However, little is known about how to design *patient-centered* technology support to promote the interpretation of test results on the part of the patients and, in turn, improve patient-provider communication [16].

Furthermore, individual patient characteristics (eg, age and sex) and medical contexts (eg, health issues and chronic conditions) are different from each other. Accordingly, the interpretation of a similar laboratory value may differ on the basis of the health condition of a patient. For example, the standard reference range of a laboratory test may not be applicable to older adults with chronic conditions [10]. Prior work has found that people often provided personal health information (eg, laboratory test results, age, medical history, and lifestyle) when posting questions on health forums to receive personalized recommendations from their peers [23,24]. From this perspective, providing information support tailored to the medical contexts of the patients could be very useful for them to determine whether results are worrisome and what might be appropriate for them to do.

In this study, we began our inquiry by asking the following research questions: (1) How to design patient-facing interfaces or tools to improve comprehension of laboratory test results for lay patients with average health literacy? (2) What system features are deemed useful (or not useful)? (3) What kinds of concerns or barriers do patients have regarding such patient-facing applications? To that end, we conducted a mixed methods, user-centered research focused on designing and evaluating an interactive prototype for communicating laboratory test results in a way that can support understanding and informed decision-making for everyone from different medical contexts and with different characteristics. More specifically, through user studies, we identified the information and technology needs of the patients related to viewing and interpreting laboratory test results. The user needs assessment informed the design of the system prototype. Finally, we conducted a pilot user evaluation with 8 patients to obtain feedback on individual design features and informational support in the system. Through this user-centered, multicomponent design exploration, we make the following contributions to the field of health informatics:

1. Design concepts for an interactive system to support patient-centered communication and comprehension of laboratory test results, as well as discussions between patients and health care providers.
2. Design implications for informative, trustworthy, and empathy-driven technology support in the context of communicating health data to lay patients.

Related Work

Patient-Centered Communication of Laboratory Test Results

The communication of clinical information, such as laboratory test results, has historically taken place during face-to-face clinical encounters. However, previous literature has pointed out various drawbacks with regard to relying solely on in-person discussions of clinical information. For instance, patients often have difficulties contacting their physicians and, thus, are not able to receive timely explanations of their laboratory test results [25]. Even during clinical visits, various barriers may hinder effective communication between patients and their physicians and, consequently, the questions the patients have are sometimes left unanswered [26-28].

Advances in patient-facing technologies, such as patient portals, enable patients to directly access laboratory test results and other personal clinical data outside of the clinical environment. The benefits of increasing the access of the patients to their data are numerous, such as enhancing patient-centered medical care [6], improving patient engagement in decision-making [1,4,5], and empowering patients to play an active role in their health care management [29]. However, the use of patient portals to review laboratory test results among patients remains limited [30]. The reason for this is multifaceted. For instance, the current interfaces of patient portals mostly present laboratory test results to patients in a tabular format, similar to the format seen by clinicians, making it challenging for patients to make sense of them [11,13,14]. Furthermore, patients with limited health literacy and numeracy find it hard to understand complex health concepts and make meaningful use of their laboratory test results (eg, determining whether they should be concerned and take action immediately) [11,31]. These challenges highlight the importance of addressing the needs and preferences of the patients when designing technology support for communicating laboratory test results. As indicated by prior work, failing to involve patients in the design process of patient-facing applications might lead to issues with technology adoption and usability [32]. To this end, our research takes a user-centered approach to explore design concepts and considerations while taking into account the informational and technological needs of the patients.

Technology for Supporting Comprehension of Laboratory Test Results

Given the complex nature of clinical data, seminal research has attempted to design health information technologies to improve people's use of clinical data. Hong et al [14] designed a system prototype to support patients, families, and health care providers in collaboratively reviewing radiology imaging data during in-person clinical visits [14]. Similarly, Arnold et al [33] developed a radiology patient portal interface to provide explanations of medical terms in lay language to help patients understand how to review radiology images. These studies have demonstrated novel techniques for supporting patient-centered communication of complex clinical data.

Despite the critical role of laboratory test results in diagnosing and screening for diseases, it remains largely unexplored how technology should be designed to support their comprehension outside of clinical settings, when the informational support that usually takes place during in-person patient-provider communication is absent. A notable exception is the study conducted by Nystrom et al [16], who designed and evaluated a new laboratory test result interface for patient portals, consisting of visual ranges of laboratory values and nontechnical descriptions of the tests. These features were deemed useful because they accounted for the needs of the patients [16]. However, one limitation of this study is that they did not address how to help patients understand the connections between their medical context and test results, and the necessary support and actions after receiving these test results. Our study bridges this important research gap.

Methods

This study consists of multiple components: user studies, an iterative prototype design, and a pilot evaluation study. All studies were approved by the Pace University Institutional Review Board.

User Studies and Prototype Design

To understand how to better support the interpretation of laboratory test results on the part of the patients through novel patient-facing technology (research question 1), we first conducted a web-based survey with 203 participants and a set of semistructured interviews with 13 patients in 2019. The user studies focused on the confusion and faced challenges of the patients pertaining to the interpretation of laboratory test results and on the informational and technological needs of the patients for better comprehension of test results. All interviews were audiotaped with the permission of the participants. Detailed information about the methodology of the user studies was reported in our previous publication [34].

The research team then used the results of the user studies to inform the design of a software prototype supporting patient-centered communication of laboratory test results. The prototype was designed in an iterative manner—after creating a design version, the researchers shared it with a small group of interview participants (n=3) for quick feedback and design improvements. This process lasted from January to June 2020.

User Evaluation

Following the prototype design, we conducted a pilot evaluation study through which we obtained responses from patients regarding individual design features and the information presented in the prototype. This evaluation study helped us answer research questions 2 and 3. We recruited 8 participants who had recently used patient portals to review laboratory test results. The demographic information is summarized in Table 1. All evaluation sessions were conducted remotely via Zoom between July and August 2020. Each session lasted 60 to 90 minutes. The consent form and a short demographic questionnaire were sent to the participants before the scheduled session. During each session, we first informed the participants about the purpose of our study and confirmed their consent to take part in it and be audio-recorded. A weblink to the prototype was sent to them so that they could explore the prototype system during the study session.

Table 1. Characteristics of the participants in the user evaluation study.

Participant ID	Sex	Age (years)	Frequency of reviewing laboratory test results[a]	Health literacy[b]
P1	Male	26-49	2-5 times	4
P2	Male	18-25	2-5 times	3
P3	Male	50-64	Once	5
P4	Female	50-64	Did not remember	5
P5	Female	18-25	Once	4
P6	Female	26-49	2-5 times	5
P7	Female	26-49	6-10 times	4
P8	Male	26-49	Once	3

[a]The number of times a participant used a patient portal to review their laboratory test results over the previous 6 months.

[b]Health literacy was self-reported by the participants on a scale of 1 to 5 (*1* denoted low literacy and *5* denoted high literacy).

We started with a demonstration of the system prototype to explain the features of the system and design rationale for each feature, as well as how patients could leverage these features to interpret their results. The participants were then encouraged to use the prototype on their own to learn more about the system. During this process, they were asked to *think aloud* [35] by

reporting their general perceptions of the system, anything that confused or surprised them, or anything that did not fully meet their expectations or needs. Since the goal of the evaluation was to obtain feedback on system features, we asked the participants to provide responses regarding whether each feature was useful rather than the nuanced design details (eg, color scheme and font size).

Once the system demonstration was completed, the researchers conducted a follow-up semistructured interview with each participant to further inquire about their experience and perceptions of using the system prototype. In particular, we sought to obtain user feedback regarding individual features, visual design, presented information, and usefulness of the system prototype. The session was concluded by administering a System Usability Scale (SUS) questionnaire—a 10-item attitude Likert scale for subjective assessments of system usability [36]. The purpose of administering the SUS questionnaire was 2-fold: (1) to assess the usability of the system prototype and (2) to collect baseline data to measure improvements for future prototype refinements. Each session was audiotaped, including verbal comments and questions from the participants about the prototype.

Data Analysis

The audio recordings of the evaluations were transcribed verbatim, and the transcripts were imported into NVivo (version 12; QSR International) for qualitative analysis. Two researchers (ZZ and LK) followed an iterative, inductive coding method [37] to analyze the transcripts and met regularly to discuss and refine codes until no new codes emerged. In the second round of analysis, coded data were grouped under themes using affinity diagrams [38]. Themes and subthemes were discussed iteratively among the researchers until a consensus was reached.

Results

In this section, we first present how the results of the user studies informed the design of a prototype application supporting patient-centered communication of laboratory test results. We then report the findings of the evaluation study.

Prototype Design for Supporting the Comprehension of Laboratory Test Results

Summary of the Results of User Studies

In this section, we briefly summarize the principal findings of the user studies to contextualize our following descriptions of how the user studies informed the system design. The detailed results of the user studies were reported elsewhere [34].

Confusion in Reviewing Laboratory Test Results

We found that there were various sources of confusion for patients regarding their test results. For example, the test report used medical jargon excessively, which is not comprehensible for lay individuals. In addition, patients were confused about how to interpret their results. In particular, when abnormal results were received, patients could not determine how serious they were. Finally, patients found it challenging to make sense of the results and their implications for their overall health care, especially when the explanations of the physicians were lacking.

Information Needs

We found that patients needed different types of information to address their confusion, including both general and personalized information. More specifically, general information needs were related to the medical terminology, reference range, and diagnostic abilities of a specific test. In contrast, many participants emphasized the importance of receiving personalized information on the basis of their medical context. First, they desired to understand the implications and causes of such abnormal test results, as well as how serious they were and if immediate action was needed. Second, there was a demand for more information about treatment options, including medications and medical procedures available for treating the medical conditions indicated by the abnormal results. Finally, a few participants wished to be informed about what actions to take next.

Technology Needs

With respect to what kind of technology support could aid their comprehension, patients emphasized that technologies should be designed for patient interpretation. In particular, the system should be user-friendly and accessible for marginalized groups (eg, older adults) to minimize disparities in the use of health technology [39]. Other approaches that were deemed useful included (1) visualizing historical results, (2) using lay terms to communicate the nature of the results, (3) including a health encyclopedia to explain medical terms, and (4) leveraging artificial intelligence to provide more personalized medical information on the basis of the medical conditions of individual patients.

Design Goals

On the basis of the results of the user studies [34], we established a list of design goals.

Facilitating Result Comprehension Using Graphical Representations and Clear Takeaway Messages

Currently, patient portals present laboratory test results in 2 main ways: dichotomously (normal vs abnormal) or through numerical values. Even though the standard reference range for each test is usually provided, patients still have difficulties understanding the seriousness of their results and whether differences between a test result and the standard reference range are significant [10,34,40]. On the basis of the literature, providing clear, plain language indications [19] and graphical representations [22,41,42] makes laboratory test results easier to review and interpret. For example, visual aids can be used to illustrate whether a result is beyond a clinically worrisome threshold. Our prototype design experimented with these cues.

Enabling Annotation of Test Reports

Patients often have a variety of questions about the different aspects of laboratory test reports. We believe that it is critical to enable patients to identify what section of the results they wish to look into or discuss further with clinicians [14]. To that end, we sought to provide annotation tools to allow patients to highlight and annotate certain medical terms and content in the application.

Clarifying Medical Jargon Using Nontechnical Verbiage

Owing to gaps in the medical domain knowledge of laypersons, the participants cited difficulties understanding medical jargon as a major barrier to the effective interpretation of test results. The literature highlights the importance of using patient-friendly language to facilitate health information–seeking and understanding, as well as informed decision-making, on the part of the patients [43-45]. Therefore, in our design, we used concise, nontechnical verbiage to describe and explain medical concepts.

Providing Pertinent and Reliable Information Tailored to the Medical Contexts of Individual Patients

When patients attempted to make sense of their results, they often turned to the internet to seek further information [23]. However, information on the web is sometimes either too general or misleading. For example, standard reference ranges of laboratory tests are not applicable to some patients with chronic conditions [10]. Our user studies indicated that patients wanted not only general information (eg, reference range and medical terms) but also personalized information (eg, treatment options, prognosis, and what to do or ask next) situated within the medical context of the patient to make sense of the normality and indications of the results. As such, our prototype was designed to present reputable and relevant information resources to help patients make sense of their results [46].

Prototyping System Features

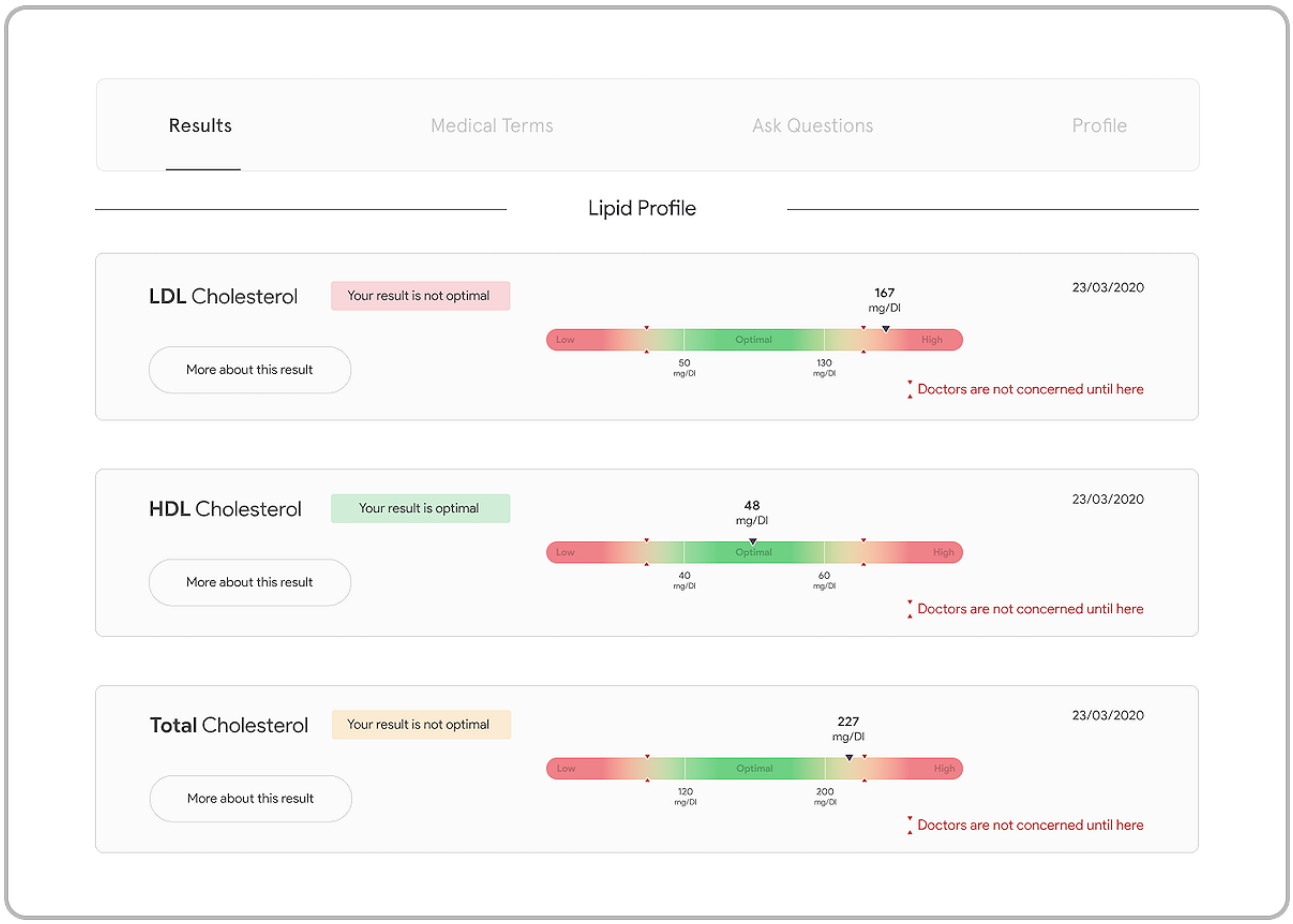
Procedure

We created a system prototype on the basis of our design goals. Because many adult patients are very likely familiar with the lipid profile (a group of laboratory blood tests on a patient to identify various levels of fat substances. Specific tests include total cholesterol, low-density cholesterol, and high-density cholesterol)—a commonly performed laboratory test among adults for many screening and diagnostic procedures—we chose to base our prototype design on this laboratory test. We used the Figma prototyping software (macOS version) to create the design. With high-fidelity animations and page transitions, this prototype enabled the user to explore how the application functions and learn each feature in an interactive manner. An HTML version of the prototype was generated for future use (eg, sent to users for feedback). In the subsequent section, we describe the main features of the prototype and how each one supports the review of the test reports.

Result Presentation

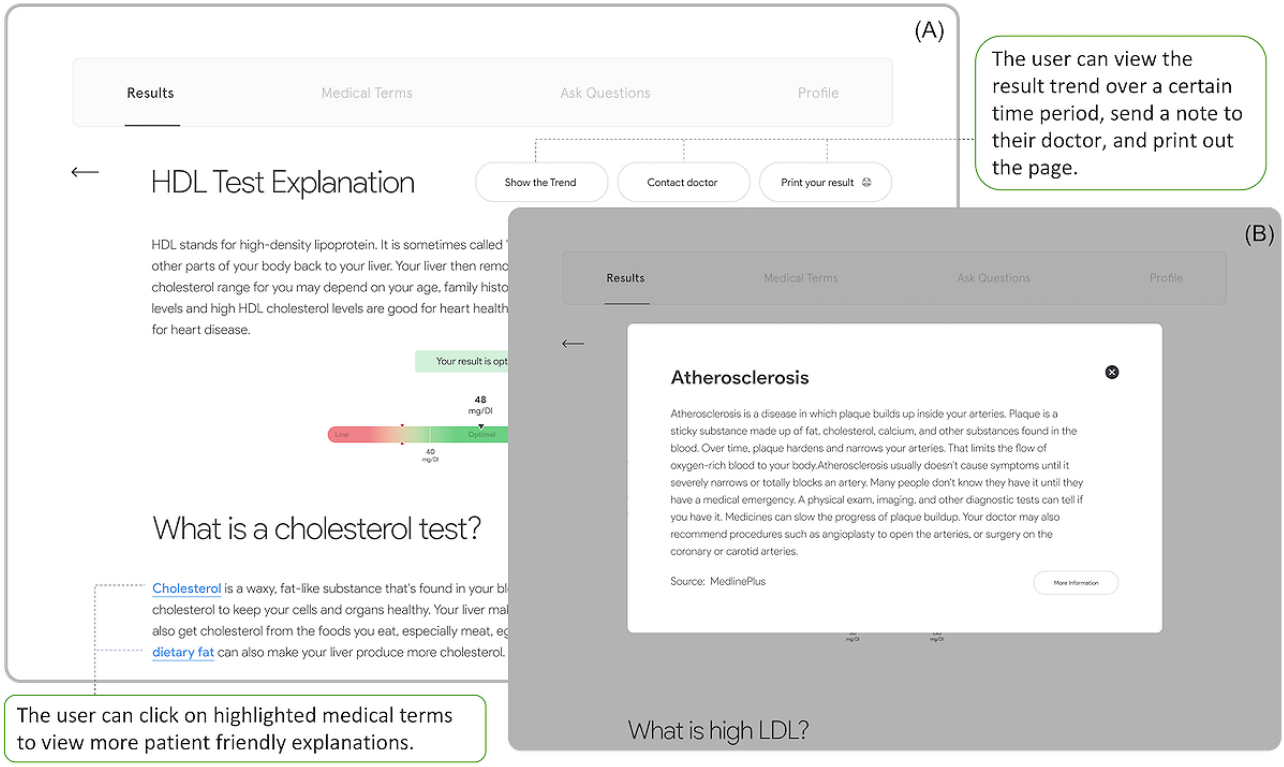
We used graphical representations, meaningful plain language, and takeaway messages to construct a new presentation interface for laboratory test results. As shown in Figure 1, gradients of 3 colors (red, yellow, and green) were used in conjunction with words (eg, *high*, *low*, and *optimal*) and takeaway messages (eg, *your result is good*) to provide an intuitive view of the normality of each test value. To further enhance the patients' understanding of the borderline values that were slightly outside of the normal range, we included a pair of red arrows and a side note (*Doctors are not concerned until here*) on each result chart to indicate at what point outside of the standard reference range the results become clinically concerning [21]. We believe that these visual aids and takeaway messages can help patients better understand the nature of their results. For example, the sample low-density cholesterol value is beyond the *clinically concerning* point, whereas the out-of-range total cholesterol value is still within the safe threshold. To further distinguish the urgencies, the takeaway message for low-density cholesterol is in red, whereas the total cholesterol takeaway message is in yellow.

Figure 1. New result presentation interface. HDL: high-density lipoprotein; LDL: low-density lipoprotein.



In addition, the user was provided with an option to learn more about each laboratory test by clicking the *More about this test* button, which directed the user to a new page containing a brief definition, detailed explanations, and a visualization of historical laboratory values (Figure 2A). The brief definition explained what a specific laboratory test was testing. Users could also find more general information related to how to deal with abnormal results (eg, *What causes high cholesterol?*). The visualization of historical laboratory values presented the trend of test results over a customizable time period (eg, 6 months or a year). This visualization could help identify the level of variation between a new result and previous results.

Figure 2. (A) Generic information about a laboratory test. (B) Explanations of a medical term. LDL: low-density lipoprotein.

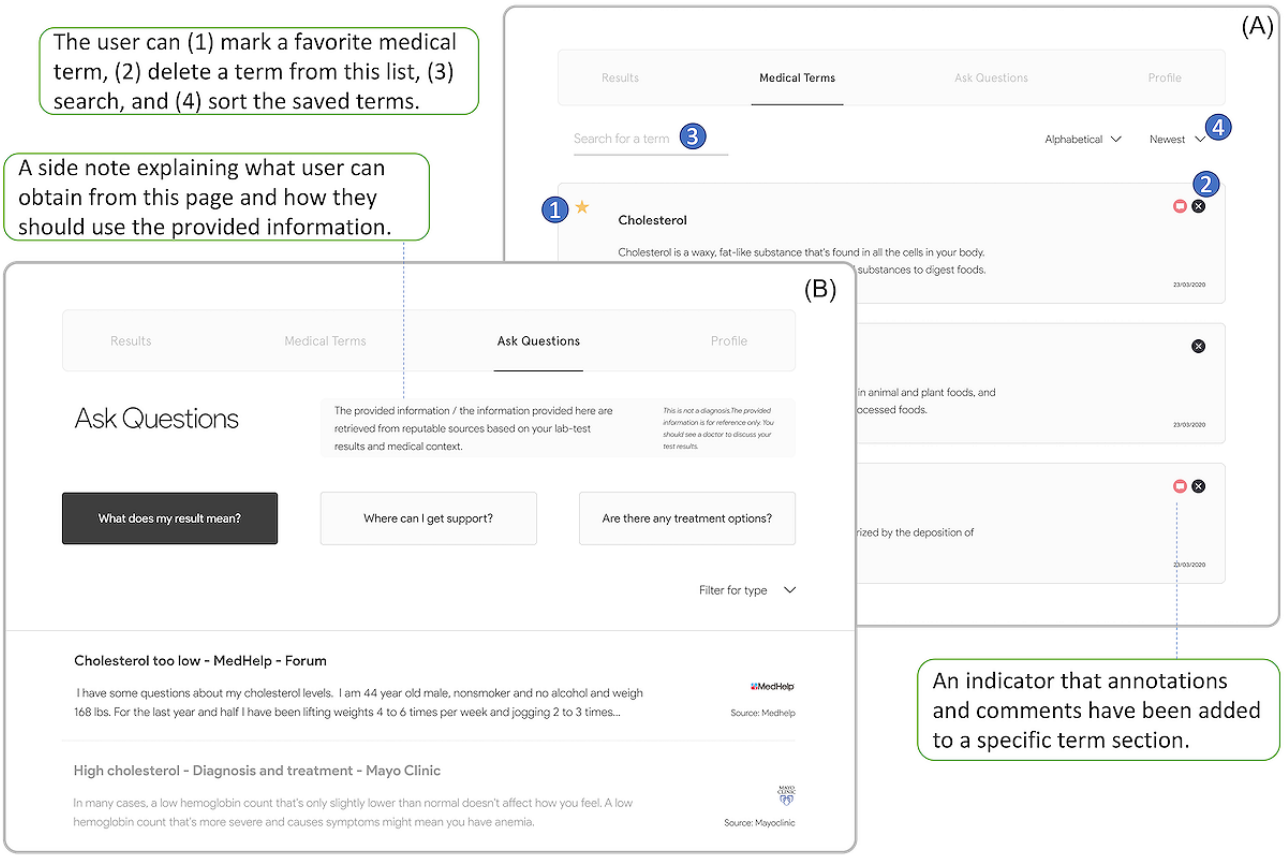


Medical Terms

Professional medical terms in the text were highlighted in blue and made clickable for the patient to learn more (Figure 2A). When a linked term was clicked, a short and concise definition was displayed in a pop-up window (Figure 2B), where users could choose to view more detailed, patient-friendly explanations retrieved from MedlinePlus (a web-based health information resource and service for patients provided by the National Library of Medicine). The sources of these explanations were noted.

In addition, every medical term the patient clicked was automatically saved to a separate page, called *Medical Terms* (Figure 3A), for future use. The user could favorite, annotate (eg, add a comment), or simply delete each saved term. All comments and annotations could be stored on the interface for later viewing by patients or clinicians (an indicator appeared if comments had been added to a specific term section). Favorited terms were always displayed at the top, followed by other saved terms. All terms added to this page could be easily searched or sorted.

Figure 3. (A) Medical Terms page. (B) Ask Questions page.



Ask Questions

The *Ask Questions* section was designed to provide *personalized* information support tailored to the demographics, medical context, and laboratory test results of the patient. On the basis of the user study results, we decided to include 3 commonly asked questions in the interface: "What does my result mean?" "Where can I get support?" "Are there any treatment options?" (Figure 3B). Information related to these 3 questions could be retrieved in real time from reputable web sources, such as health care organizations (eg, Mayo Clinic), online health care forums (eg, MedHelp), and medical literature databases (eg, PubMed). The original source of the presented information was also provided, enabling patients to decide whether they wanted to trust and use that information. In light of medical ethics [47], users were advised that the information was for reference only and that the application was not designed as a diagnostic tool.

Since the primary goal of this study was to explore and evaluate design concepts and considerations rather than implementing a fully functional system, we decided to hardcode information on the basis of a hypothetical patient case. For example, we conducted a search on MedHelp using hypothetical laboratory values and patient age and sex and retrieved a post containing similar information. We then added that post to the "What does my result mean?" section and provided a link to the original post.

Other Features

The application provided an annotation tool on all pages, allowing patients to add comments and highlight texts. This tool was expected to facilitate reading of and reflection on the results and discussions between patients and physicians during in-person clinical visits.

The application also provided a straightforward onboarding process through which a patient could easily import their laboratory test report from the patient portal of the health care provider. The medical information (eg, chronic conditions) and characteristics (eg, age and sex) of the patient could also be imported and then stored in the *Profile* page.

Pilot System Prototype Evaluation

In this section, we describe the pilot evaluation sessions with 8 patients and report their responses to individual design features and the information presented in the prototype.

Overall Perception

The participants uniformly acknowledged the significant improvement of the prototype over the current patient portal design and expressed excitement about using the new application to review and interpret their results:

> *It looks extremely robust and I am basing that on my comparisons of what I have now, which is nowhere near, doesn't come anywhere near this functionality. So it's a vast improvement over what I have now.* [P3]

The average SUS score was 95 (out of 100), signifying that the prototype was perceived as very user-friendly. In the subsequent section, we describe the detailed feedback on each of the main features of the system.

Result Presentation

All participants had a positive response to the new presentation interface of the test results. One primary improvement was the use of visual cues in combination with takeaway messages to intuitively communicate the nature of the results. As the participants stated:

> *I think the colors are really helpful and that having that indicator on the side correspond with the colors saying, "your whatever is good" and then you have "low" or "high" corresponds with the color. It's all very easy to read. I really like the whole package, everything is very clean.* [P7]

> *I really like the "doctors are not concerned until here" thing, because sometimes my numbers are low. It's not extremely low, but it's still low, and I'm not sure if I should be concerned or not.* [P2]

The participants also appreciated the ability to learn more about each test:

> *You have the very obvious button to get more information that you might need. I think for that use, it's perfect for what you have on there. Not too much, not too little. I think it's pretty good.* [P8]

When asked about what aspects of the new result presentation interface could be further improved, the feedback was mostly focused on the wordings. For example, 2 participants mentioned that some words in the takeaway messages (eg, *not optimal*) could trigger unnecessary anxiety. They suggested that the communication of abnormal laboratory test results to patients should take their emotions into consideration:

> *You don't want to panic anyone. You don't want anyone to read anything that they're going to immediately have to be on the phone with their doctor, right? [. . .] Maybe just keeping that in mind when you have the results page set up. [. . .] Having some sort of caveat that these blood results are just results they're not diagnose. [. . .] You need to be empathetic when giving bad results.* [P6]

Medical Terms

All participants agreed that automatically highlighting medical terms in the text and providing detailed explanations with patient-friendly language was very useful, and this convenience of getting to know medical terms could improve the experience of reviewing test results:

> *It's probably one of the most useful things. Because when I am interpreting my personal lab results, I'd be back and forth going to a search engine to look up and find out what it is. That would save me the trouble of doing that if I could just click on it.* [P3]

With regard to automatically saving every clicked or viewed medical term to a dedicated page (*Medical Terms*), we received diverse feedback. Most participants (n=6) stated that they found it very useful for future use, but some concerns were raised. For example, one participant mentioned that she might not be interested in using and reviewing the *Medical Terms* page:

> *I do think it is useful but I don't know if it's totally necessary like the Results section is. I could see me not really using the Medical Terms section as much as the Results or the Ask Questions section.* [P7]

Another participant pointed out that users may not be aware that the clicked medical term was automatically saved to a new page for future use. As such, they suggested redesigning this user flow to increase awareness of this potentially useful feature.

Ask Questions

All participants considered the *Ask Questions* section one of the most useful features because it provided pertinent information tailored to the medical context of individual patients:

> *I think it is actually potentially one of the coolest parts of the whole thing for every different condition that a person could have and for all the different things a person could have on their chart. If it all comes up with really good, concise information for that person to take further, I could see this being incredibly useful.* [P8]

When asked if including only 3 prescribed questions was sufficient, the participants had positive responses:

> *I would say that those three questions cover what would effectively be my general concern.* [P1]

However, they also suggested adding "what to ask during a clinical visit" to this page:

> *It might be cool to have a section or information about what you should ask your doctor.* [P7]

Even though we specified the sources of the provided information, some participants (n=3) still had concerns regarding information trustworthiness because each person has their own trusted information sources. For example, the information collected from well-known health organizations was deemed more reliable than the information from peer-to-peer online forums:

> *If you took me to a peer to peer forum, I wouldn't trust it since they are not always the most reliable thing and there's a lot of misinformation there. I think if you would like to use such things, you might want to always have a caveat with something like "please be aware of peer to peer responses." You get my point? [. . .] You should link more to credible information sources, such as JAMA or Lancet.* [P4]

Annotation

The perceived usefulness of the annotation tool varied among the participants. Some (n=4) acknowledged that it was useful to take notes for future use, remarking that:

> *I'm going to look for the things in here that I feel like are the most important and I would definitely annotate*

> *those so that I can scroll quickly back through it if I needed a refresher.* [P1]

However, others (n=4) stated that they very likely would not use this tool, as one participant explained:

> *Personally, I think its usefulness is marginal. I probably wouldn't use it. I'm probably going to come here [the application] and I'm going to read what I want to read and pick up what I want to pick up. That's not saying that someone else might not want to highlight some things, maybe they would. [. . .] But I doubt seriously that personally I would use it. I mean, it's kind of cool, but I don't think I'd really miss it if it wasn't there.* [P3]

Another participant echoed this statement and further explained that she was a "paper person," so she was not comfortable using the digital annotation tool:

> *Personally, I don't really think that I would use it. I would print this out and highlight stuff to show or read when I see my doctor.* [P2]

During follow-up discussions on how to make the annotation tools more useful, half the participants (including one who previously had a positive attitude in this regard) suggested consolidating all the annotations in one page for easy reading, management, and retrieval:

> *It would be really cool to have a section sort of like the "Medical Terms" where you can save questions to ask your doctor. So you can just go to your doctor's office for the follow-up and pull up the app and ask the questions right from the app.* [P7]

Another participant shared a similar idea and expressed the desire to be able to share all the notes and questions with his primary physician before the clinical visit:

> *We all know that we only have 15 minutes with our doctors, and they prefer we email them most of the time now. If you could export the notes, and your questions about what you saw here into one document, and then get them to my doctor through a messenger, that would be great.* [P4]

General Feedback

Interestingly, the participants did not express any privacy concerns about the use of their personal health information to generate personalized information support:

> *Any realm of health care is so highly regulated. I understand that there are breaches, but I don't personally have that concern. So I feel like if people have already used patient portals then they're not necessarily going to have those concerns and the ones that are super concerned aren't going to use any sort of patient portal.* [P1]

In fact, they stated that they understood why their personal health data were needed, and they were willing to provide more information to receive tailored support.

Regarding the aspects of the application that could be improved, the participants provided some interesting suggestions. For example, one participant raised that people with different levels of health literacy may need different types of information and, as such, the system should be designed to first assess their health knowledge and then present information tailored to their literacy level:

> *So, if I'm new to the medical condition, it is going to be different than if it's something that I've had for a while and I understand some of the medical terminologies, right? So it may be useful to have a variety of information sources and suggestions for the user with different health literacy backgrounds. You can have very basic things for a new patient but for someone that has had the medical condition for a while, you are going to want maybe some academic articles with most recent research results or that kind of thing.* [P6]

In addition, some participants recommended adding explanations to the presented information to avoid potential confusion or misunderstandings:

> *The one thing that you might want to have a little caveat somewhere, [stating] that those ranges are created based on what is normal for a certain percentage of a population so if you're outside of range doesn't necessarily mean that you will show signs and symptoms. [. . .] You can have a little note as to where these ranges come from so when people look at their result and see high or low, where they are on the spectrum, they are having an idea of where those come from and if the normal ranges apply to him or her.* [P6]

Finally, it was deemed useful to provide both a web and mobile version of the system to expand its use scenarios, for example, it is easier to use mobile version during in-person clinical visits, whereas many users preferred a web-based version at home:

> *Having a mobile application with you while you're at your doctor visit might help to more quickly review the results and get your doctor on board.* [P1]

Discussion

Principal Findings

In this study, we designed and evaluated an interactive system prototype to support patients in making sense of their laboratory test results. The system design was informed by user studies, the details of which were reported elsewhere [34]. The system prototype consisted of several novel, interactive features: (1) using graphical representations and clear takeaway messages to convey the nature of the results, (2) enabling users to annotate laboratory test reports, (3) clarifying medical jargon using nontechnical verbiage and allowing users to interact with the medical terms (eg, saving, favoriting, or sorting), and (4) providing pertinent and reliable information tailored to the medical contexts of individual patients. Through a pilot user evaluation, potential users uniformly acknowledged the significant improvement of the system over current patient portals in communicating laboratory test results to patients. They noted that the new system could facilitate their

interpretation of laboratory test results by promoting self-education regarding different aspects of laboratory test reports.

In addition, we identified two other use cases in which the new system could play a significant role. One was helping patients be better prepared for clinical consultations. Although it varies by health care setting, there is evidence suggesting that clinicians are spending less time with each patient and are often unable to fully address their questions [48]. Even more concerning is that some clinicians frequently interrupt patients, making it difficult for them to have their questions answered during clinical encounters [49]. Our application can better prepare patients for clinical consultations so that they can use their time with physicians effectively. For example, both general and personalized information provided in the application could help the patient research different aspects of the laboratory report and, in turn, contemplate and devise questions related to the results before the consultation.

The other use case regarded facilitating patient-provider discussions during clinical visits. For example, the application allows patients to annotate and document comments or questions while viewing the results. These annotations can be used later during clinical visits to facilitate patient-provider discussions. Although not mentioned by the participants, we envision that these annotation tools can also be used by patients to take notes during doctor visits. Patients can then use the notes to support memory recall of the information discussed during the consultation. Our future work will iteratively design and evaluate interactive features, such as annotation tools, that can further enhance in situ review of laboratory test data with clinicians.

Design Implications

Informative Technology Support

It is common to see patients feeling helpless when they receive laboratory test results because current patient portals provide limited support for them to assimilate the information and make informed decisions. Our user studies highlighted the importance of providing additional information that patients could read more about to begin conducting their own research [34]. For example, even though 2.5 mU/L in the first trimester and 3.0 mU/L in the second and third trimesters are considered the standard reference ranges for the thyroid-stimulating hormone during pregnancy, it was pointed out that these cut-offs are too low and may lead to overtreatment [50]. It might be useful to provide this information to pregnant patients to raise their awareness.

Furthermore, the participants appreciated the ability to receive information tailored to their medical context so that they could learn what the results meant, what they could do next, and where to obtain support. A key example was individualizing the standard reference range—given that a typical standard reference range for laboratory test results is developed on the basis of a large, healthy population, it may not be applicable to certain populations, such as pregnant women, older adults, or people with comorbidities [10]. In this case, it is useful to individualize the frame of reference by allowing custom reference ranges. Given that many known or unknown factors could affect what reference ranges might be desired for a patient, this may be more appropriately done by their health care provider, in which case they should be granted access to the system as well.

Promoting Trust

One major feature of our application is the provision of additional information tailored to the medical context of the patient. Despite its perceived usefulness, some participants expressed concerns about the credibility and trustworthiness of the information provided. In this prototype, we provided the original source of the information to alleviate this issue. In-depth analyses of trusted information sources revealed that some people preferred information provided by well-known health care organizations, whereas a few others wished to read scientific literature published by reputable medical journals. From this perspective, it is necessary to tailor the delivery of additional information to the individual differences of the patients. One way to accomplish this is by allowing patients to customize their trusted and preferred information sources within the application to individualize information delivery. In addition, it would be helpful to provide a mechanism for patients to rate the usefulness and trustworthiness of each information source, and the ratings can be used to create a curated list of reputable information sources.

Future development of personalized support in this context is expected to rely on advanced machine learning techniques, which can take patient characteristics, medical contexts, and laboratory values as inputs to find and retrieve relevant and up-to-date medical evidence. The literature on the perceived trustworthiness of machine learning and artificial intelligence technologies has suggested presenting a variety of system-related information to the users to help them better understand how the recommendations of the system are generated and then determine whether it is appropriate to trust them [51]. For example, Lee and See [52] suggested presenting system reliability (eg, how reliable and accurate the system is) and logic and reasoning (eg, how the system operates), as well as the types of information that the system leverages (or excludes) to generate recommendations. In future work, aligning with these suggestions, we will examine whether providing more appropriate explanations for how the system generates medical advice could be of any help in promoting acceptance on the part of the patients and fostering trust in the application.

Empathy-Driven Design

In the medical field, the types of clinical data that should be shared with patients and how the data should be communicated have been the subject of debate among informatics researchers [53]. Some clinicians are concerned about providing patients with direct access to clinical data, such as laboratory test results, before clinical consultations [54]. The primary reason is that clinicians worry that patients, especially those who receive abnormal test results, are vulnerable to anxiety and frustration when reviewing the results without the presence of clinicians [12]. In contrast, patients are increasingly interested in having direct access to their test results, regardless of their normality [4,18]. The key, then, is how to communicate test results and their potential implications to patients while accounting for their emotional needs. Prior work has suggested that patient-facing

applications that communicate sensitive information (eg, abnormal laboratory test results) should be designed with *empathy*. However, such needs have been largely unfulfilled by current patient-facing systems, even though the empathy-driven design has been gaining momentum for many years [55].

In our prototype design, we took into account this design consideration. For example, we followed suggestions by prior work [21] to provide visual aids to indicate at what point outside of the standard range the results become clinically worrisome with the aim of lowering patient distress if they receive slightly out-of-range test values. We also provided a list of web-based sources in the *Ask Questions* section, where patients could seek and receive emotional support. In future work, we will examine additional features and strategies that can mitigate emotional stress. For example, as one of the participants suggested, it might be useful to add a simple caveat in plain language to address the concerns of the users (eg, "this standard range is only applicable to 80% of people, many factors, such as chronic health conditions and characteristics, could impact what might be the appropriate frame of reference for you"). We will also work closely with communication specialists and medical professionals to synthesize best practices and strategies for communicating abnormal results to patients and then incorporate them into the system refinement.

Conclusions and Limitations

In this paper, we described a multicomponent study focused on the design of a patient-centered system prototype to help patients interpret highly professional laboratory test results. Through user studies, we identified patient informational and technological needs specific to this domain, which were then used to inform a set of design considerations and concepts. After iterative prototyping, an initial evaluation study was conducted to obtain feedback from 8 patients, who uniformly had positive responses and acknowledged the significant improvement over existing patient portals in supporting comprehension of laboratory test results on the part of the patients. Finally, on the basis of our findings, we discussed design implications for communicating personal clinical data to lay individuals in a more informative, trustworthy, and empathetic manner.

A few limitations of this study should be noted. First, we only conducted an initial evaluation study with a small sample size (n=8). However, they all had extensive experience with laboratory tests and had reviewed laboratory test results via patient portals. All of them acknowledged the significant improvement of our application design over the current patient portals. We will use the feedback received to refine the system prototype. Second, in the pilot evaluation, our focus was mainly on the responses of the patients to each design feature or concept with respect to whether or not the design or system feature was useful. As such, we primarily collected subjective assessment data. We did not evaluate the knowledge gain of the patients in using our system or the extent to which the system could help patients make sense of each result and inform their decision-making and subsequent actions. These aspects will be assessed in future user evaluation sessions. Third, we used a hypothetical scenario with only one type of laboratory test (lipid profile). This limitation could affect the generalizability of the results of our study. As we continue to refine the prototype and address its shortcomings, we plan to explore how these design concepts and considerations play out in different medical scenarios, as well as whether they are able to support the interpretation of different types of laboratory tests and can be used by different patient populations. Finally, because this was a user-centered design exploration study to investigate design considerations for supporting patient comprehension of laboratory test results, we did not fully implement the system. This is a common practice in user-centered design research [56]. For example, to prototype the feature of presenting relevant information tailored to the context of the patient, we manually searched information on the web based on the hypothetical patient case, collectively determined which information source to use, and then hardcoded the information into the application. In future work, we will implement the system and integrate it with existing patient portals once the system design is finalized.

Acknowledgments

This study was supported by the Pace University Scholarly Research Award, the National Institute on Aging of the National Institutes of Health under award R21AG061431, and the Florida State University-University of Florida Clinical and Translational Science Award funded by the National Center for Advancing Translational Sciences under award UL1TR001427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Ahern DK, Woods SS, Lightowler MC, Finley SW, Houston TK. Promise of and potential for patient-facing technologies to enable meaningful use. Am J Prev Med 2011 May;40(5 Suppl 2):162-172. [doi: 10.1016/j.amepre.2011.01.005] [Medline: 21521591]
2. Berlin L. Communicating results of all radiologic examinations directly to patients: has the time come? Am J Roentgenol 2007 Dec;189(6):1275-1282. [doi: 10.2214/ajr.07.2740]

3. Peacock S, Reddy A, Leveille S, Walker J, Payne T, Oster NJ, et al. Patient portals and personal health information online: perception, access, and use by US adults. J Am Med Inform Assoc 2017 Apr 01;24(e1):173-177 [FREE Full text] [doi: 10.1093/jamia/ocw095] [Medline: 27413120]
4. Baldwin DM, Quintela J, Duclos C, Staton EW, Pace WD. Patient preferences for notification of normal laboratory test results: a report from the ASIPS Collaborative. BMC Fam Pract 2005 Mar 08;6(1):11 [FREE Full text] [doi: 10.1186/1471-2296-6-11] [Medline: 15755328]
5. Johnson AJ, Easterling D, Nelson R, Chen MY, Frankel RM. Access to radiologic reports via a patient portal: clinical simulations to investigate patient preferences. J Am Coll Radiol 2012 Apr;9(4):256-263. [doi: 10.1016/j.jacr.2011.12.023] [Medline: 22469376]
6. O'Kane MJ, Lopez B. Explaining laboratory test results to patients: what the clinician needs to know. Br Med J 2015 Dec 03;351:h5552. [doi: 10.1136/bmj.h5552] [Medline: 26634382]
7. Walker J, Meltsner M, Delbanco T. US experience with doctors and patients sharing clinical notes. Br Med J 2015 Feb 10;350:g7785. [doi: 10.1136/bmj.g7785] [Medline: 25670182]
8. Ross SE, Moore LA, Earnest MA, Wittevrongel L, Lin C. Providing a web-based online medical record with electronic communication capabilities to patients with congestive heart failure: randomized trial. J Med Internet Res 2004 May 14;6(2):e12 [FREE Full text] [doi: 10.2196/jmir.6.2.e12] [Medline: 15249261]
9. O'Kane M, Freedman D, Zikmund-Fisher BJ. Can patients use test results effectively if they have direct access? Br Med J 2015 Feb 11;350:h673. [doi: 10.1136/bmj.h673] [Medline: 25673132]
10. Witteman HB, Zikmund-Fisher BJ. Communicating laboratory results to patients and families. Clin Chem Lab Med 2019 Feb 25;57(3):359-364. [doi: 10.1515/cclm-2018-0634] [Medline: 30407910]
11. Zikmund-Fisher BJ, Exe NL, Witteman HO. Numeracy and literacy independently predict patients' ability to identify out-of-range test results. J Med Internet Res 2014 Aug 08;16(8):e187 [FREE Full text] [doi: 10.2196/jmir.3241] [Medline: 25135688]
12. Giardina T, Baldwin J, Nystrom D, Sittig DH, Singh H. Patient perceptions of receiving test results via online portals: a mixed-methods study. J Am Med Inform Assoc 2018 Apr 01;25(4):440-446 [FREE Full text] [doi: 10.1093/jamia/ocx140] [Medline: 29240899]
13. Alpert JM, Krist AH, Aycock RA, Kreps GL. Applying multiple methods to comprehensively evaluate a patient portal's effectiveness to convey information to patients. J Med Internet Res 2016 May 17;18(5):e112 [FREE Full text] [doi: 10.2196/jmir.5451] [Medline: 27188953]
14. Hong M, Feustel C, Agnihotri M, Silverman M, Simoneaux S, Wilcox L. Supporting families in reviewing and communicating about radiology imaging studies. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17: CHI Conference on Human Factors in Computing Systems; May 6 - 11, 2017; Denver Colorado USA p. 5245-5256. [doi: 10.1145/3025453.3025754]
15. Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. Med Care Res Rev 2007 Apr 19;64(2):169-190. [doi: 10.1177/1077558707060020301] [Medline: 17406019]
16. Nystrom DT, Singh H, Baldwin J, Sittig DF, Giardina TD. Methods for Patient-Centered Interface Design of Test Result Display in Online Portals. EGEMS (Wash DC) 2018 Jun 26;6(1):15 [FREE Full text] [doi: 10.5334/egems.255] [Medline: 30094287]
17. Giardina T, Modi V, Parrish D, Singh H. The patient portal and abnormal test results: an exploratory study of patient experiences. Patient Exp J 2015 Apr 29;2(1):148-154. [doi: 10.35680/2372-0247.1055]
18. Pillemer F, Price RA, Paone S, Martich GD, Albert S, Haidari L, et al. Direct release of test results to patients increases patient engagement and utilization of care. PLoS One 2016 Jun 23;11(6):e0154743 [FREE Full text] [doi: 10.1371/journal.pone.0154743] [Medline: 27337092]
19. Elder NC, Barney K. “But what does it mean for me?” Primary care patients’ communication preferences for test results notification. Jt Comm J Qual Patient Saf 2012 Apr;38(4):168. [doi: 10.1016/s1553-7250(12)38022-7]
20. Torsvik T, Lillebo B, Mikkelsen G. Presentation of clinical laboratory results: an experimental comparison of four visualization techniques. J Am Med Inform Assoc 2013 Mar 01;20(2):325-331 [FREE Full text] [doi: 10.1136/amiajnl-2012-001147] [Medline: 23043123]
21. Zikmund-Fisher BJ, Scherer AM, Witteman HO, Solomon JB, Exe NL, Fagerlin A. Effect of harm anchors in visual displays of test results on patient perceptions of urgency about near-normal values: experimental study. J Med Internet Res 2018 Mar 26;20(3):e98 [FREE Full text] [doi: 10.2196/jmir.8889] [Medline: 29581088]
22. Zikmund-Fisher B, Scherer A, Witteman H, Solomon J, Exe N, Tarini BA, et al. Graphics help patients distinguish between urgent and non-urgent deviations in laboratory test results. J Am Med Inform Assoc 2017 May 01;24(3):520-528 [FREE Full text] [doi: 10.1093/jamia/ocw169] [Medline: 28040686]
23. Zhang Z, Lu Y, Kou Y, Wu DT, Huh-Yoo J, He Z. Understanding patient information needs about their clinical laboratory results: a study of social q and site. Stud Health Technol Inform 2019 Aug 21;264:1403-1407 [FREE Full text] [doi: 10.3233/SHTI190458] [Medline: 31438157]
24. Zhang Y. Contextualizing consumer health information searching: an analysis of questions in a social Q and A community. In: Proceedings of the 1st ACM International Health Informatics Symposium. 2010 Presented at: IHI '10: ACM International

Health Informatics Symposium; November 11 - 12, 2010; Arlington Virginia USA p. 210-219. [doi: 10.1145/1882992.1883023]
25. Bodenheimer T, Pham HH. Primary care: current problems and proposed solutions. Health Aff (Millwood) 2010 May;29(5):799-805. [doi: 10.1377/hlthaff.2010.0026] [Medline: 20439864]
26. Ma X, Gui X, Fan J, Zhao M, Chen Y, Zheng K. Professional medical advice at your fingertips. Proc ACM Hum-Comput Interact 2018 Nov;2(CSCW):1-22. [doi: 10.1145/3274385]
27. Chen Y, Cheng K, Tang C, Siek KJ, Bardram JE. Is my doctor listening to me?: impact of health it systems on patient-provider interaction. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; April 27 - May 2, 2013; Paris France p. 2419-2426. [doi: 10.1145/2468356.2468791]
28. Park S, Chen YS, Raj S. Beyond health literacy: supporting patient-provider communication during an emergency visit. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17: Computer Supported Cooperative Work and Social Computing; February 25 - March 1, 2017; Portland Oregon USA p. 2179-2192. [doi: 10.1145/2998181.2998357]
29. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc 2011 Jul 01;18(4):515-522 [FREE Full text] [doi: 10.1136/amiajnl-2011-000105] [Medline: 21672914]
30. Reynolds T, Ali N, McGregor E, O'Brien T, Longhurst C, Rosenberg A, et al. Understanding patient questions about their medical records in an online health forum: opportunity for patient portal design. AMIA Annu Symp Proc 2017;2017:1468-1477 [FREE Full text] [Medline: 29854216]
31. McCray AT. Promoting health literacy. J Am Med Inform Asso 2004 Nov 23;12(2):152-163. [doi: 10.1197/jamia.m1687]
32. Wilcox L, Patel R, Chen Y, Shachak A. Human factors in computing systems: focus on patient-centered health communication at the ACM SIGCHI conference. Patient Educ Couns 2013 Dec;93(3):532-534. [doi: 10.1016/j.pec.2013.09.017] [Medline: 24184039]
33. Arnold CW, McNamara M, El-Saden S, Chen S, Taira RK, Bui AA. Imaging informatics for consumer health: towards a radiology patient portal. J Am Med Inform Assoc 2013 Nov 01;20(6):1028-1036 [FREE Full text] [doi: 10.1136/amiajnl-2012-001457] [Medline: 23739614]
34. Zhang Z, Citardi D, Xing A, Luo X, Lu YZ, He Z. Patient challenges and needs in comprehending laboratory test results: mixed methods study. J Med Internet Res 2020 Dec 07;22(12):e18725 [FREE Full text] [doi: 10.2196/18725] [Medline: 33284117]
35. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. Qual Health Res 2016 Jul 01;3(4):430-441. [doi: 10.1177/104973239300300403]
36. Brooke J. SUS: A 'Quick and Dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, Florida, United States: CRC Press; 1996:189-194.
37. Boyatzis R. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, CA: Sage Publications Inc; 1998:1-200.
38. Holtzblatt K, Beyer H. Contextual design: defining customer-centered systems. In: Proceedings of the Human Factors in Computing Systems Conference. 1998 Presented at: CHI98: ACM Conference on Human Factors and Computing Systems; April 18 - 23, 1998; Los Angeles California USA p. 149-150. [doi: 10.1145/286498.286629]
39. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]
40. Solomon J, Scherer A, Exe N, Witteman H, Fagerlin AB, Zikmund-Fisher BJ. Is this good or bad? Redesigning visual displays of medical test results in patient portals to provide context and meaning. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Presented at: CHI'16: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose California USA p. 2314-2320. [doi: 10.1145/2851581.2892523]
41. Haggstrom D, Saleem J, Russ A, Jones J, Russell SN, Chumbler NR. Lessons learned from usability testing of the VA's personal health record. J Am Med Inform Assoc 2011 Dec;18 Suppl 1:13-17 [FREE Full text] [doi: 10.1136/amiajnl-2010-000082] [Medline: 21984604]
42. Sittig D, Murphy D, Smith M, Russo E, Wright AH, Singh H. Graphical display of diagnostic test results in electronic health records: a comparison of 8 systems. J Am Med Inform Assoc 2015 Jul;22(4):900-904 [FREE Full text] [doi: 10.1093/jamia/ocv013] [Medline: 25792704]
43. Keselman A, Tse T, Crowell J, Browne A, Ngo L, Zeng Q. Assessing consumer health vocabulary familiarity: an exploratory study. J Med Internet Res 2007 Mar 14;9(1):e5 [FREE Full text] [doi: 10.2196/jmir.9.1.e5] [Medline: 17478414]
44. Zeng QT, Tse T. Exploring and developing consumer health vocabularies. J Am Med Inform Asso 2006 Jan 01;13(1):24-29. [doi: 10.1197/jamia.m1761]
45. Zeng Q, Tse T, Crowell J, Divita G, Roth LA, Browne AC. Identifying consumer-friendly display (CFD) names for health concepts. AMIA Annu Symp Proc 2005:859-863 [FREE Full text] [Medline: 16779162]
46. Keselman A, Smith CA. A classification of errors in lay comprehension of medical documents. J Biomed Inform 2012 Dec;45(6):1151-1163 [FREE Full text] [doi: 10.1016/j.jbi.2012.07.012] [Medline: 22925723]

47. Doshi-Velez F, Kortz M, Budish R, Bavitz C, Gershman SJ, O'Brien D, et al. Accountability of AI under the law: the role of explanation. SSRN J 2020:1-21. [doi: 10.2139/ssrn.3064761]
48. Mechanic D, McAlpine DD, Rosenthal M. Are patients' office visits with physicians getting shorter? N Engl J Med 2001 Jan 18;344(3):198-204. [doi: 10.1056/nejm200101183440307]
49. Rhoades DR, McFarland K, Finch WA, Johnson AO. Speaking and interruptions during primary care office visits. Fam Med 2001;33(7):528-532. [Medline: 11456245]
50. Korevaar TIM. The upper limit for TSH during pregnancy: why we should stop using fixed limits of 2.5 or 3.0 mU/l. Thyroid Res 2018 May 21;11(1):5 [FREE Full text] [doi: 10.1186/s13044-018-0048-7] [Medline: 29942352]
51. Vorm EA. Assessing the value of transparency in recommender systems: an end-user perspective. In: Proceedings of the 5th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems at ACM Conference on Recommender Systems (RecSys 2018). 2018 Presented at: 5th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems at ACM Conference on Recommender Systems (RecSys 2018); October 2018; Vancouver, Canada URL: https://www.researchgate.net/publication/328134005_Assessing_the_Value_of_Transparency_in_Recommender_Systems_An_End-User_Perspective
52. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors 2004;46(1):50-80. [doi: 10.1518/hfes.46.1.50.30392]
53. Halamka JD, Mandl KD, Tang PC. Early experiences with personal health records. J Am Med Inform Asso 2008 Jan 01;15(1):1-7. [doi: 10.1197/jamia.m2562]
54. Giardina T, Singh H. Should patients get direct access to their laboratory test results? An answer with many questions. J Am Med Assoc 2011 Dec 14;306(22):2502-2503. [doi: 10.1001/jama.2011.1797] [Medline: 22122864]
55. Choe EK, Duarte ME, Suh H, Pratt W, Kientz JA. Communicating bad news: insights for the design of consumer health technologies. JMIR Hum Factors 2019 May 17;6(2):e8885 [FREE Full text] [doi: 10.2196/humanfactors.8885] [Medline: 31102374]
56. Dopp A, Parisi K, Munson SA, Lyon AR. A glossary of user-centered design strategies for implementation experts. Transl Behav Med 2019 Nov 25;9(6):1057-1064. [doi: 10.1093/tbm/iby119] [Medline: 30535343]

Abbreviations

SUS: System Usability Scale

Edited by A Kushniruk; submitted 09.12.20; peer-reviewed by D He, H Trachtman; comments to author 12.01.21; revised version received 19.01.21; accepted 11.09.21; published 25.11.21.

Please cite as:
Zhang Z, Kmoth L, Luo X, He Z
User-Centered System Design for Communicating Clinical Laboratory Test Results: Design and Evaluation Study
JMIR Hum Factors 2021;8(4):e26017
URL: https://humanfactors.jmir.org/2021/4/e26017
doi:10.2196/26017
PMID:34842529

Original Paper

Evaluating An Automated Compounding Workflow Software for Safety and Efficiency: Implementation Study

Ülle Helena Meren[1*], MPharm; James Waterson[2*], MMedEd

[1]East Tallinn Central Hospital, Tallinn, Estonia

[2]Medical Affairs, Medication Management Solutions, Becton Dickinson Ltd, Dubai, United Arab Emirates

[*]all authors contributed equally

Corresponding Author:
James Waterson, MMedEd
Medical Affairs
Medication Management Solutions
Becton Dickinson Ltd
11F Blue Bay Tower
Business Bay
Dubai, 52279
United Arab Emirates
Phone: 971 566035154
Email: james.waterson@bd.com

Abstract

Background: The forms of automation available to the oncology pharmacy range from compounding robotic solutions through to combination workflow software, which can scale-up to cover the entire workflow from prescribing to administration. A solution that offers entire workflow management for oncology is desirable because (in terms of cytotoxic delivery of a regimen to a patient) the chain that starts with prescription and the assay of the patient's laboratory results and ends with administration has multiple potential safety gaps and choke points.

Objective: The aim of this study was to show how incremental change to a core compounding workflow software solution has helped an organization meet goals of improved patient safety; increasing the number of oncology treatments; improving documentation; and improving communication between oncologists, pharmacists, and nurses. We also aimed to illustrate how using this technology flow beyond the pharmacy has extended medication safety to the patient's bedside through the deployment of a connected solution for confirming and documenting right patient–right medication transactions.

Methods: A compounding workflow software solution was introduced for both preparation and documentation, with pharmacist verification of the order, gravimetric checks, and step-by-step on-screen instructions displayed in the work area for the technician. The software supported the technician during compounding by proposing the required drug vial size, diluents, and consumables. Out-of-tolerance concentrations were auto-alerted via an integrated gravimetric scale. A patient-medication label was created. Integration was undertaken between a prescribing module and the compounding module to reduce the risk of transcription errors. The deployment of wireless-connected handheld barcode scanners was then made to allow nurses to use the patient-medication label on each compounded product and to scan patient identification bands to ensure right patient–right prescription.

Results: Despite an increase in compounding, with a growth of 12% per annum and no increase in pharmacy headcount, we doubled our output to 14,000 medications per annum through the application of the compounding solution. The use of a handheld barcode scanning device for nurses reduced the time for medication administration from ≈6 minutes per item to 41 seconds, with a mean average saving of 5 minutes and 19 seconds per item. When calculated against our throughput of 14,000 items per annum (current production rate via pharmacy), this gives a saving of 3 hours and 24 minutes of nursing time per day, equivalent to 0.425 full-time nurses per annum.

Conclusions: The addition of prescribing, compounding, and administration software solutions to our oncology medication chain has increased detection and decreased the risk of error at each stage of the process. The double-checks that the system has built in by virtue of its own systems and through the flow of control of drugs and dosages from physician to pharmacist to nurse allow it to integrate fully with our human systems of risk management.

(JMIR Hum Factors 2021;8(4):e29180) doi:10.2196/29180

KEYWORDS

compounding; medication safety; positive patient identification; gravimetric; automation; closed loop

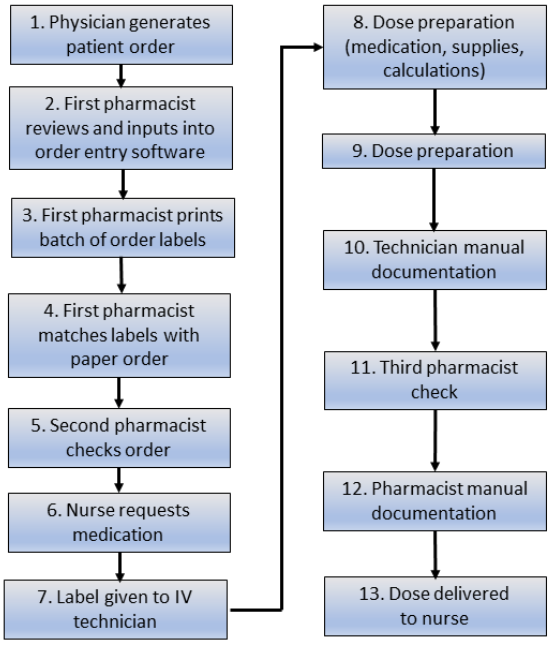
Introduction

Background

A systematic review of the literature from 2020 related to automated compounding technology and workflow solutions for the preparation of chemotherapy concluded that "implementation of chemotherapy compounding automation solutions may reduce compounding errors and reduce costs; however, this is highly variable depending on the form of automation" [1].

In terms of scaling up compounding, managing the entire workflow for oncology therapy management, from prescription via compounding and through to administration, is the most logical solution to the increasing demands that have been seen in both oncology inpatients and outpatients and which will continue to increase in the near future. In fact, one estimate suggested that by 2040, globally, the number of patients requiring at least first-line chemotherapy each year would have increased from a 2018 baseline by 53% (ie, from 9.8 million to 15 million individuals) [2]. An entire workflow management for cytotoxic prescription, production, and administration is desirable because in terms of cytotoxic delivery of a regimen to a patient the chain that starts with prescription and the assay of the patient's laboratory results and ends with administration has multiple potential safety gaps and choke points (Figure 1).

Figure 1. A semimanual intravenous oncology medication chain with safety gaps and productivity choke points. Adapted from Reece et al [3].



The possible errors and safety flaws of the aforesaid compounding process in the pharmacy begin with possible errors of transcription at Step 2. Order entry software may have dose-limiting features, which may reduce the risk of dose transcription error, but which may miss area under the curve (AUC) dose reductions in the original order. Without integration into prescription software the risk of "simple" lookalike–soundalike transcription errors also exists. The pressure for delivering multiple patient doses and the fact that the pharmacist is often only present at key stages for technician checks gives Step 3, the pre-emptive printing of a batch of order labels, the potential for causing mix-ups of patient-product labels with incorrect product labeling at Steps 7-10. Besides, under this system the third pharmacist check at Step 11 is of marginal value in a batching process as used ampoules, carriage fluids, and labeled final patient products may be verified but with no guarantee of accuracy nor correct patient-product matching in any unwitnessed steps (highly likely to be Steps 8-10) The issuing of a label checked against the prescription at Step 5 actually *precedes* the physical creation of products that then have these labels applied. This does not follow a logical failure mode effect analysis (FMEA) process [4], where steps involving risk must take place before any final verification checks and the issuing of a label with a unique preparation and patient identification (ID) number. The second pharmacist check at Step 5 similarly precedes the actual creation of patient products and is therefore, in terms of FMEA, redundant.

In terms of the above system's capacity for maintaining consistent patient-product supply or responding to increased demand there are also several choke points.

Transcription from an unintegrated computerized provider order entry (CPOE) system into the pharmacy compounding system (Step 2) requires the work of 1 pharmacist and replaces clerical work for more useful or appropriate tasks. Steps 3-5 are similarly manual and are perhaps retained as they give a sense of security that the process is under the control of well-qualified pharmacists. Step 11 requires the physical presence of the pharmacist in the compounding clean room, an area that is geographically separate from the main pharmacy suite in our

unit. With distractions from other ongoing tasks, both supervisory and specialist, it can be difficult to coordinate between technicians and pharmacists to achieve rapid checks, and this is a major choke point, because it not only slows dispensing of completed patient products, but also causes backlogs as other products cannot be compounded until unchecked patient products are released. Step 12 can delay release of patient products to the nursing unit, as in a compounding system that is unintegrated with administration this is the last point at which the compounding unit has visibility over the medication; therefore, manual documentation must take place before patient products can be transported for administration. This is another clerical action that takes the pharmacist away from higher-value tasks.

Variations of doses or dose adjustments close to the time of therapy due to late-phase AUC adjustments based on the patient's laboratory results can also cause wastage or require recompounding in rigid systems that require longer loading and setting times.

There is of course a requirement for accuracy of compounding beyond those relating to AUC alterations. Compounding errors may be of magnitudes significant enough to cause direct patient harm, and the problem persists despite advances in workflow management and technological assistance in the compounding space. A recent survey of both pharmacists and pharmacy technicians found that 74% of all respondents were aware of at least one compounding error in the past 12 months, including those discovered in the pharmacy and after dispensing, with incorrect dose or concentration being the most cited error (58%) of all of those discovered [5].

A 2016 study of the implementation of an oncology compounding workflow software solution showed the gravimetric component of the solution catching 797 deviations from acceptable tolerances (<4%) before injection into the final administration intravenous (IV) bag (11,874 preparations in total). Catches at this stage of the workflow are significant, with the possibility of reworking the dose, and the study noted that no deviations were detected at the final weight verification step, with the correct amount of drug accurately injected into the final administration IV bag [3].

The 2020 survey [5] also identified other errors beyond final dosing/concentration, with incorrect base solution identified by 51% of respondents and incorrect reconstitution of a drug in terms of volume or diluent stated by 36%. Furthermore, only 52% of respondents reported that "it is *always* easy to identify with certainty which (and how many) drugs, diluents, and volumes were used when verifying the preparation of each Compounded Sterile Procedure." It is notable in this respect that the 2016 study identified how a "no-software" FMEA system detected only 1 wrong diluent event, whereas the compounding workflow software solution when integrated into a new FMEA system detected 52 such events. The FMEA risk priority score (severity score × probability score × detectability score) for "wrong fluid selected" in the 2016 study dropped from 567 to 108 after the software application was implemented [3], chiefly because of a significant drop in the risk of detection failure.

The difficulty of error detection during high-risk processes such as compounding and administration of IV chemotherapy is an issue where technology can undoubtedly assist. The emerging evidence related to medication administration at the bedside, a part of the medication delivery chain where currently it has been suggested that at least 38% of all medication errors occur [6], is that introducing a final barcode scan–based check of "right patient–right medication" using patient ID labels and barcoded medications that include the patient's medical record number and the "order string" pertaining to the patient's particular regimen and prescription may reduce the overall error rate by as much as 3:1 [7,8]. Without the presence of "nonhuman" confirmatory processes in place it has been suggested that the detectability of administration error falls as low as 2%, and that of dispensing/compounding error versus prescription error falls to 34% [6].

Human FMEA systems for compounding emphasize the double-check of each stage of the process with pharmacist oversight of technicians. It is likely, however, that pharmacists, given their workload, the reduced numbers of qualified staff available against a backdrop of increasing demand on health care services, and the closed nature of the sterile compounding unit, can only be present for "key stages" of the compounding process. For example, the key parts of the compounding (diluent, medication vials, closed system transfer devices, final administration IV bag, and recipe) may be shown to the pharmacist as a "guarantee" of correct constituents for compounding, but given that lookalike–soundalike errors remain prevalent in pharmacies (estimated at 25.9% of all errors) [9] and that in hurried checks the "4-eyes" process may only reinforce error rather than avert it [10,11], this is far from optimal. To this end, a compounding process that has other monitoring processes outside of the assumed presence and infallibility of a second human check is desirable. A compounding workflow software solution that allows for electronic verification and documentation of each preparation from end-to-end with ideally image recognition and capture that can document workarounds, such as "supermarket-style" scanning of the same ampoule several times for multiple vial usage, can give this level of real-world evidence. Equally, rejected patient-medication scans at the bedside could assist us in identifying a little more of the iceberg of this error, as currently the other established methods are very much retrospective because they are based on chart review [12] or reliant on self-reporting, with all its attendant issues [13].

Auto-documentation of medication administration directly into the patient's record is certainly superior to manual completion of the medication record, because such documentation is commonly delayed or inaccurate as clinicians attend to emergent situations or distractions [14]. Once clinicians return to their documentation after a patient care event, such as medication administration, they often transcribe from memory. Having a secondary nonhuman confirmation of patient and medication matching via barcode scanning would be of huge value for audit even in systems that are transitioning between electronic prescribing and paper documentation of administration to be able to compare scan-library data with manual chart entries.

Gravimetric systems are an integral part of some compounding workflow software solutions, and their main function has been to ensure dosing is within tolerance. A large-scale European study that ran over 4 years [15] showed how a gravimetric system detected a 7.89% error rate (nearly 60,000 errors) for compounded doses outside of tolerance in a total of 759,060 doses of antineoplastic drugs. Over 10% deviations were seen in a mean of 2.25% (range 0.49%-5.04%) and over 20% deviations were seen in a mean of 0.71% (range 0.21%-1.27%) of compounded medications.

Estonia faces the same pressures seen in other Organisation for Economic Co-operation and Development (OECD) countries: an increasing number of patients with cancer, an acute lack of medical personnel, and increasing restrictions on budgets.

Before the implementation of a gravimetric compounding workflow software solution in the pharmacy serving the Oncology Department of the East Tallinn Central Hospital, nurses prepared all cytotoxic medications on the oncology day care unit. The process was entirely unautomated and undertaken under a biosafety hood using closed-system transfer devices. The same nurses who prepared the medications also administered them. The amount of time spent preparing medications detracted from time spent on patient care and there was no comprehensive documentation of the medication regimen. The workload was becoming untenable by 2012 due to increasing complexity of treatments and increasing numbers of patients.

The organization therefore set itself 4 initial goals:

- Increase the number of oncology treatments.
- Improve patient safety.
- Improve documentation.
- Improve communication between the oncologist, the pharmacist, and the nurse.

Objectives

The overall objective was to show how incremental change to a core compounding workflow software solution has helped the organization meet the above goals, has released nursing time, and has acted as a catalyst to extend medication safety beyond the compounding of medications in the central pharmacy and to the patient's bedside, where a further technology enhancement has also improved efficiency, documentation, and confidence in the closing of the medication chain with right patient–right medication transactions being verified and documented.

Methods

Materials

The following materials were utilized for the solution implemented: a compounding workflow software solution (BD Cato); a prescribing workflow software solution (BD Cato Prescribe); a closed-system transfer device (BD PhaSeal); a barcode medication administration (BCMA) software suite (BD Cato ReadyMed); a handheld user interface on the bedside; and a BCMA-enabled bedside handheld device (Zebra TC56 Mobile Computer).

Study Design

The establishment of compounding, prescribing, and administration software and solutions was incremental. Each component was, however, essentially undertaken under conditions of a ceteris paribus pre–post study design except the increasing volume of chemotherapy required and delivered, as staffing, physical space, transportation methods, and communication channels between the compounding unit and nursing units remained unchanged. We undertook a qualitative review of the changes made with staff and within our team structure supported by a retrospective quantitative review of compounding production capabilities over 8 years, and an ex-ante and post-ante quantitative review of time required for nurses to confirm patient ID and product match prior to administration of compounded products with the introduction of BCMA capabilities over a period of 4 weeks.

The process of selection of hardware and software for the move from nurse-led to pharmacy-centered compounding was undertaken in the light of the studies above. The final selection was a compounding workflow software solution, along with the continuance of a CSTD and an existing Class A isolator unit. The CSTD was known to nursing staff, so its continuance of use was logical, as it would reduce workflow change.

In the first build, a configuration for the compounding workflow software solution to manage both preparation and documentation was initiated. The suite had initial pharmacist verification of the order, gravimetric checks, and on-screen instructions displayed on the monitor inside the work area for the technician. The user interface requires minimal interactions during the compounding process. The technician simply follows step-by-step instructions. For example, the interface will propose a list of items that includes drug, diluents, and the consumables required to prepare the dose. Automatic dosage calculations are undertaken by the software according to the prescribed regimen. The use of preset regimens was extended as much as possible to increase standardization and reduce divergence from workflow. Scanning of individual components ensures a recipe match for all items including the final administration IV bag. The system also carries hard stops for dosing out of tolerance as per the recommendations of the studies above [6,15]. Besides, the system only delivers a patient-medication label after all the steps of compounding are successfully completed, which makes it ideal for building into an FMEA process, as the steps involving risk are all before the final verification checks and the issuing of the label with a unique preparation and patient ID number, which reduces the risk of administering the medication to the wrong patient. This was fundamental to our later project to allow for right patient–right medication checks at the bedside.

The software records all cancelled mixes, tolerance limit breaches, incorrect item scans, and the resolutions of alerts by the user. Each distinct group of events from first alert to resolution is recorded. Date–time stamps are applied to all of these alerts. Data are continuously collected from the compounding logs and stored locally on an MS-SQL database inside the hospital firewall. Documentation in the central pharmacy of compounding statistics for each preparation is

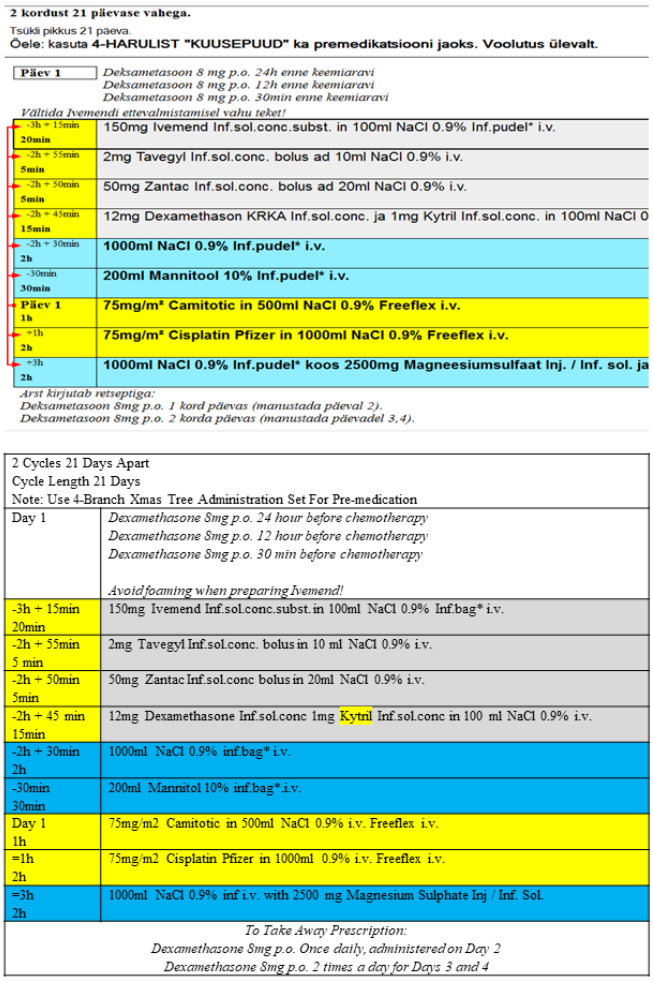
aggregated and assists with forecasting, and each preparation is date–time stamped.

All pharmacy staff are aware of this ongoing collection and analysis of near-miss events. This is important if we want to get as close as possible to "normal behavior" with our data. As with all observational and self-reporting studies, the Hawthorne effect remains a very real danger. The advantage with "passive" data collection, such as that gathered by the software, is that users will not alter their behavior as they might during a time-limited study. This philosophy of consent through thorough understanding of the nature of data collection was later extended to the nursing trial of handheld medication-patient barcode scanning devices.

Integration between the CPOE system and the pharmacy compounding software suite modules was the next intervention undertaken with the addition of the prescribing module, chiefly to reduce the risk of transcription errors. A 2015 study [16] described how near-miss transcription error (NMTE) reporting rates varied between an institution's formal reporting system built on traditional lines of self-reporting of near-miss and identified-incident reporting and an adapted NMTE reporting mechanism utilizing an error queue within the institution's order imaging software. The NMTE system described in the study was similar to the video capture component of our compounding solution but in fact did not have an integration between prescription and compounding. However, it is a useful guide to the number of NMTEs that might be avoided through application of a system that eliminates the need for transcribing and removes a risk element from the process. In this study the data collection spanned 92 days during which time about 460,000 medication orders were processed. In total, 1563 NMTEs were reported using the transcription error queue imaging software (0.34%), while only 12 errors were reported (detected) via the formal reporting mechanism (0.003%).

The prescribing software also includes hard stops for dosing and automation of calculations for AUC dosing. Physician prescriptions are received electronically into the pharmacy for verification, and if approved, for push communication to the compounding unit. The module also gives access for prescribers to computer-based standardized protocols, which reduce the number of nonstandard regimens requiring creation, can make for a faster regimen build, and automatically calculate doses and creates preparation guidance. A typical regimen is shown in Figure 2.

Figure 2. A patient-specific cycle from a regimen as presented to the prescriber, compounder, and nurse via prescription software (with English translation).



2 kordust 21 päevase vahega.
Tsükli pikkus 21 päeva.
Öele: kasuta 4-HARULIST "KUUSEPUUD" ka premedikatsiooni jaoks. Voolutus ülevalt.

Päev 1	*Deksametasoon 8 mg p.o. 24h enne keemiaravi* *Deksametasoon 8 mg p.o. 12h enne keemiaravi* *Deksametasoon 8 mg p.o. 30min enne keemiaravi*
	Vältida Ivemendi ettevalmistamisel vahu teket!
-3h + 15min **20min**	150mg Ivemend Inf.sol.conc.subst. in 100ml NaCl 0.9% Inf.pudel* i.v.
-2h + 55min **5min**	2mg Tavegyl Inf.sol.conc. bolus ad 10ml NaCl 0.9% i.v.
-2h + 50min **5min**	50mg Zantac Inf.sol.conc. bolus ad 20ml NaCl 0.9% i.v.
-2h + 45min **15min**	12mg Dexamethason KRKA Inf.sol.conc. ja 1mg Kytril Inf.sol.conc. in 100ml NaCl 0
-2h + 30min **2h**	**1000ml NaCl 0.9% Inf.pudel* i.v.**
-30min **30min**	**200ml Mannitool 10% Inf.pudel* i.v.**
Päev 1 **1h**	**75mg/m² Camitotic in 500ml NaCl 0.9% Freeflex i.v.**
+1h **2h**	**75mg/m² Cisplatin Pfizer in 1000ml NaCl 0.9% Freeflex i.v.**
+3h **2h**	**1000ml NaCl 0.9% Inf.pudel* koos 2500mg Magneesiumsulfaat Inj. / Inf. sol. ja**
	Arst kirjutab retseptiga: *Deksametasoon 8mg p.o. 1 kord päevas (manustada päeval 2).* *Deksametasoon 8mg p.o. 2 korda päevas (manustada päevadel 3,4).*

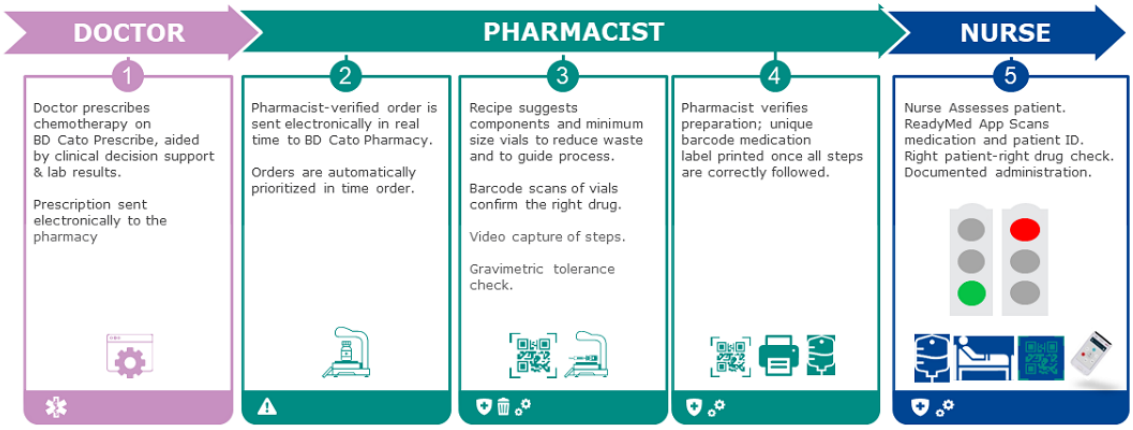
2 Cycles 21 Days Apart Cycle Length 21 Days Note: Use 4-Branch Xmas Tree Administration Set For Pre-medication	
Day 1	*Dexamethasone 8mg p.o. 24 hour before chemotherapy* *Dexamethasone 8mg p.o. 12 hour before chemotherapy* *Dexamethasone 8mg p.o. 30 min before chemotherapy* *Avoid foaming when preparing Ivemend!*
-3h + 15min 20min	150mg Ivemend Inf.sol.conc.subst.in 100ml NaCl 0.9% Inf.bag* i.v.
-2h + 55min 5 min	2mg Tavegyl Inf.sol.conc. bolus in 10 ml NaCl 0.9% i.v.
-2h + 50min 5min	50mg Zantac Inf.sol.conc bolus in 20ml NaCl 0.9% i.v.
-2h + 45 min 15min	12mg Dexamethasone Inf.sol.conc 1mg Kytril Inf.sol.conc in 100 ml NaCl 0.9% i.v.
-2h + 30min 2h	1000ml NaCl 0.9% inf.bag* i.v.
-30min 30min	200ml Mannitol 10% inf.bag* i.v.
Day 1 1h	75mg/m2 Camitotic in 500ml NaCl 0.9% i.v. Freeflex i.v.
=1h 2h	75mg/m2 Cisplatin Pfizer in 1000ml 0.9% i.v. Freeflex i.v.
=3h 2h	1000ml NaCl 0.9% inf i.v. with 2500 mg Magnesium Sulphate Inj / Inf. Sol.
	To Take Away Prescription: *Dexamethasone 8mg p.o. Once daily, administered on Day 2* *Dexamethasone 8mg p.o. 2 times a day for Days 3 and 4*

The compounding library itself was created by, and is updated and confirmed by, the Pharmacy and Therapeutics Committee.

A key update that requires regular review is any changes in the specific gravity of core medications as this will impact on the gravimetric check.

The postimplementation workflow (Figures 2 and 3) had changed substantially from the "classic" manual compounding unit workflow described in Figure 1.

Figure 3. Extension of the patient-medication matching solution from prescribing to the compounding unit and to the bedside via ReadyMed.



We collected data on throughput and set a key performance indicator for reducing time of preparation as a response to receiving no increased full-time employee (FTE) headcount and the need to meet the unit needs, which were forecast to increase at 10%-12% per annum. At the preimplementation stage, the compounding team in the pharmacy, using manual techniques, was averaging ≈6 minutes per preparation. The time was calculated as a mean average of the time taken from the beginning to the end for the compounding of individual 1-ingredient products, with the below steps:

1. Pharmacist reviews and inputs into the order entry software (measurement starts).
2. Pharmacist prints order labels and matches labels with prescription.
3. Second pharmacist checks labels versus order.
4. Technician takes labels and prepares dose: medication selection, supplies and diluents, calculations.
5. Technician documentation.
6. Final pharmacist check of the product.
7. Pharmacist documentation of the product being ready for dispatch to the nursing unit (measurement stopped).

No metrics for mixing by nurses before implementation were available, as their role was split between preparation and administration.

Wastage was addressed by taking advantage of reissuing options in the software, and through the activation of an advisory within the software that proposes the use of a drug vial size that will result in the least amount of waste for the prescribed dose to be compounded. Analysis of each preparation's data, and aggregation of these data, assisted us in managing and optimizing the inventory, monitoring drug wastage, and measuring productivity.

The second stage of the project involved an extension into the inpatient unit with the deployment of wireless-connected handheld barcode scanners. This allowed us to take advantage of the patient-medication label on each compounded product via the BCMA device (handheld at the bedside).

The move to BCMA was seen as a desirable part of our build for patient safety and efficiency, and the BCMA administration system we envisaged was to feed directly from the prescription module software and to obtain its products for administration from an integrated pharmacy module.

The BCMA handheld interface was initially only available in English but was intuitive enough for the launch; an Estonian language product was available later in the project. The process of scanning the patient for positive patient ID, and then scanning the product to be administered triggers a matching of patient and product information from the compounded product's barcode to the patient's ID and to the prescription via the BCMA software and at the interface of the prescribing server. There was and remains a regular process of engagement with nursing leadership and clinical educators to introduce functional changes to the workflow. Acceptance of the new process was good. The new workflow (Figure 3) has replaced a large amount of manual activity by nursing staff and should help to reduce the risk of medication errors by a substantial degree given the literature findings above. The 2 nurse or "4-eyes" check is not a common practice in our facility and is not mandated in Estonia. Prior to the software implementation, the physician printed, verified, and signed the therapy plan, and handed it over to the nurse who matched the paperwork with the product and then with the patient. The process was laborious and charts and order sheets were at risk of being mislaid and were commonly not readily at hand to be aligned and checked against each other.

For audit purposes the processing time of each patient administration can be calculated from date–time stamps on the handheld device, but before implementation data had to be gathered manually. Before and after the implementation of the ReadyMed solution, staffing remained unchanged with an average of 3.5 nurses on unit duty, 3 pharmacy technicians, and 2 pharmacists. The inpatient unit remained at an 8-bed capacity. Walking time between the nursing workstation, where initial checks of the received compounded products is performed, and the patient rooms was unchanged at 10.8 minutes per day of "travel time." By this point the pharmacy compounding unit was producing ≈14,000 cytotoxic products per annum.

The preimplementation observation was undertaken with consent from nursing staff on a daily basis, as personnel changed on each shift. It was made clear that personal performance would not be identifiable in quantitative results, although data would be continuously collected by the devices. A short timeframe (4 weeks) was deliberately applied for the BCMA ex- and post-ante review to reduce the risk of data distortion arising from increased throughput of patients and the rising number of products compounded and dispensed for administration. This period was long enough to ensure "capture" of all nursing staff during both introduction and training periods. The study type was a pre–post design in that all other factors were ceteris paribus including the staff involved (all nursing staff used the handheld scanner and had used the traditional paper-based method extensively, and all had equal amounts of training and exposure to the new system).

The mean batch of medications to be given by each nurse per shift, mean averaged over the working year both before and after the handheld scanner implementation, was 15 items (SD 2.7). Quantitative measurements of time taken to process medications in each system were therefore based on an average of total time per 15-item groupings rather than time per single-item measurement to reduce the risk of bias from one-off measurements or from possible clusters of "simple" regimens and single items, or of additionally complex regimens.

Study Procedure

The data were patient anonymized, and no personal information items such as clinician ID, hospital number, gender, name, date of birth, diagnosis, or other identifiable material were recorded for analysis.

BD Clinical Management and Global Customer Service were engaged to optimize the solution and BD Medical Affairs were requested to undertake a deeper analysis of the data. The medical affairs department of BD operates as a distinct arm outside of the commercial operations of the company.

Inclusion Criteria

All cytotoxic infusions compounded from within the oncology formulary (and therefore identifiable in terms of medication name, dose, and duration as per cycle usage over the period) were included in the study. These included weight-based and non-weight–based infusions and body surface area–based infusions.

Exclusion Criteria

Infusions that did not require compounding such as flush bags and preregimen, premixed hydration infusions that do not pass through the compounding unit were excluded from the study.

Results

Our forecasts for growth in both patient throughput and the requirement for compounded oncology medications were reasonably accurate. In fact, growth has been 12% overall, with more than 16,000 patient visits per year (outpatient, daily clinic, and inpatient short stay). Compounding production has met this increase without an increase in FTE headcount (Multimedia Appendix 1).

Despite this increased load and unchanged FTE staffing, there was an overall reduction in compounding time of 35% using the same start and endpoints applied in our measurement of preimplementation checks and compounding times (from "order enters system" through to "product available for delivery to nursing unit"). We believe there are savings in improved management of remnants, but quantifying this would be difficult without a full accounting of pre- and post-implementation ampoule usage per comparable volumes of prescriptions compounded. We do not currently have these data available.

The project using the ReadyMed handheld barcode scanning solution showed substantial nurse time savings within a relatively short period. Within 2 weeks a mean average time of 41 seconds (0.697 minutes) was required for each product–patient matching and verification of the order by the system. When calculated against our throughput of 14,000 items per annum, we saw considerable nurse-time savings, as shown in Tables 1 and 2.

For the qualitative component of the BCMA study we identified the following categorizations of statements with an incidence of above 80% (8/9, 89%) in responses after 2 weeks of use of the handheld scanning devices (Textbox 1). The overall satisfaction was measured by means of a numerical scale. Categorization by statements with responses with an incidence of greater than 80% (8/9, 89%) was undertaken. Given the small sample size, this is as much complexity for analysis as could be achieved. The nursing workforce is very small.

Table 1. Nursing time saved per day.

Statistics	Time for each daily batch (minutes per nurse)	Number of medications (each processed daily batch per nurse; N=321)
Mean (SD)	10.476 (10.666)	15.286 (7.191)
Median	7	14

Table 2. Sensitivity analysis on Q1 and Q3 and range for 14,000 compounded medications scenarios.

Parameter	Verification speed (minutes per item)	Time saving per item (minutes)	FTE[a] gain per 14,000 items PA[b] assuming 225 days of work PA per FTE	Nursing time saved per day (hours and minutes per 14,000 items)
Mean	0.697	5.303	0.425	3 hours and 24 minutes
Min	0.222	5.778	0.46	3 hours and 41 minutes
Max	1.875	4.125	0.33	2 hours and 38 minutes
Q1	0.316	5.684	0.45	3 hours and 38 minutes
Q3	1.000	5.000	0.40	2 hours and 11 minutes

[a]FTE: full-time employee.

[b]PA: per annum.

Textbox 1. Grouping and statement categorization with incidence of greater than 80% (8/9, 89%) in responses (N=9).

Perception of safety/confidence

- "Easier to identify patient"
- "Chance of medical error was less"
- "Device prompts user to identify the patient"
- "Notification if you try to administer the wrong drug"

Usability

- "Easier to follow the drug chart"
- "Saves time for documentation"

Documentation

- "Local language would increase the use"
- "With one scan it marked who administered, which drug, and at what time"

Discussion

In classic FMEA planning [4], for any high-risk activity, and particularly for those with a high risk of "low-chance or no-chance" of detection of error, the activity is broken down into a number of steps, each of which can mitigate, correct, or annul any error in the previous steps. The addition of prescribe–compound–administration software solutions to our oncology medication chain has increased detection, and decreased the risk, of error at each stage of the medication chain. What is significant in the process that we describe is that this took a considerable period (over 8 years) to reach its end points.

We could see that the issues of clerical tasks most heavily impacted the compounding unit, and given the limited budget, our 4 goals to increase production and patient safety were the most pressing demands. We focused initially on the workflow described in Figure 1. The benefits for productivity were significant with the choke points at Steps 7, 10, and 12 being removed. In this interim period, Steps 3 and 4 moved to the end of the process where they could be of real value in the FMEA approach, and Steps 8 and 9 were made safer by implementing gravimetric checks and hard stops, as well as guidance from the "recipe" screen. Our outcomes matched those of the study by Reece et al [3] in terms of no deviations being detected in the final weight verification step, a "good catch" rate of less than 4% for out-of-tolerance compounded medications before their injection into the final administration IV bag, and injection of the correct amount of drug accurately into the final administration IV bag in all compounded final products, as verified by the software's documentation of the process through video capture and recording of gravimetric data.

However, at this stage it was not possible to avoid the transcription risks in Step 2. For a considerable period the final pharmacist's check therefore had to focus on the paper prescription for validation of the final compounded product to reduce the risk of checking against an invalid transcription, though this was supported by the documentation of the compounding process through video captures and process recording.

As noted previously, it had been hard to coordinate between technicians and pharmacists to undertake key point checks. With the new system, however, technicians did not need to wait for the pharmacist's physical availability for checks, as these can be made via the process recording. This allowed for faster release of products from the clean room and the release of this choke point on throughput.

We could see from our road map that investment in the prescribing module would achieve improved patient safety given the findings of the NMTE study of 2015 [16] and it would again reduce the clerical load significantly. Step 2 of Reece et al's [3] semimanual oncology medication compounding chain and the

risk of NMTE were essentially eliminated with the addition of the prescribing to compounding modules. The elimination of the CPOE to pharmacy system transcription requirement also ended the requirement for the physical presence of the pharmacist in the compounding clean room, as orders now feed directly from the CPOE to the "recipe" screens.

Manual documentation was also eliminated at this point, but printing of the regimen was still required as the administration roadblock remained. The system, as an interim measure, still enabled us to print patient-medication orders and match attendant medication labeling against patients' case notes, and for its identification against patient IDs; however, as productivity and demand for compounded products had increased significantly (Multimedia Appendix 1), it was clear to us that administration also needed to move from a paper to electronic form to be in harmony with prescribing and compounding.

We recognize the limitations of the approaches we have undertaken. One key issue was the protracted length of time that the total study took place over, with ongoing changes to regimens and addition and deletion of medications, and the lack of preautomation data that were available to us for comparing and contrasting. Most of the key elements including staffing numbers, the physical environment, and the supply chain have essentially remained unchanged over this period, which has allowed us to accept the assumption that the changes in performance have been related to the introduction of the 3 solutions. This would not necessarily be the case in most units or facilities with fluctuating headcounts, physical unit changes, and changes in supply processes. This said, while the length of time over which the compounding solution was reviewed should reduce the risk of bias, as it is distinctly long-run data (please note the small drop-off in production in 2020 due to the COVID-19 emergency), there is a risk that the benefits of BCMA which we saw are not reflected on such a significant scale in larger units or in specialist patient subpopulations.

There are very few oncology units in Europe or globally that have a similar experience to us in the area of BCMA handheld devices for IV medication administration. It would be a natural progression of this study to a multicenter compounding study, including units with larger headcounts, and in environments with different medication delivery services to the oncology unit.

On a national level the project and its outcomes have been significant. The ability to see a patient's treatment history with a single click fits well with Estonia's vision of digital hospitals with entirely paper-free documentation. We are now documenting any side effects through the prescribing software, and we are expanding on features such as doctors and pharmacists being able to leave specific administration cautions or notes for nurses to be picked up at the patient ID and medication ID stage of administration. The next stage of the project should be to extend it across a larger hospital campus system and networking the main hospital with our partner hospitals. This expansion in the region is important as it may drive health technology vendors toward accelerating product localizations. Adapting interfaces to local language increased acceptance of the new technologies in our experience, although even with English interfaces initially being used staff were enthusiastic about the changes, and their individual learning curves were not influenced by age or general technology acceptance, but by their acknowledgement of the advantages the systems gave, particularly in terms of protecting patients. We found that the technology was quickly adopted by all staff after only a few days. The qualitative survey of nurses was encouraging in this respect with statements such as "chance of medical error was less" and "notification if you try to administer the wrong drug" being given as positives.

Studies of error in health care have found that most serious errors occur during the execution of treatment, with "performance-level failures outweighing rule-based or knowledge-based mistakes" [17]. For this reason, we are very positive about the presence of hard stops rather than advisories in the software for prescribing, compounding, and administration for both dosing and medication components and patient ID. Our staff are highly skilled and experienced but are human.

Acknowledgments

Publication and analysis support was provided by the Eastern Europe, Middle East, and Africa Medical Affairs Department of Becton, Dickinson and Company (BD), a manufacturer of a broad portfolio of health care products, which includes automated dispensing products. The company is committed to enhancing patient, clinician, and environmental safety through scholarly scientific research.

Conflicts of Interest

UM certifies that she has no financial affiliations with, or involvement in, any organizations or entities with a financial interest. JW certifies that he has no financial affiliations with, or involvement in, any organizations or entities with a financial interest, beyond his employment in the Medical Affairs Department at Becton, Dickinson and Company (BD).

Multimedia Appendix 1
Total oncology medications compounded per annum. Implementation of the compounding workflow software solution took place in late 2012.
[PNG File , 18 KB - humanfactors_v8i4e29180_app1.png]

References

1. Batson S, Mitchell SA, Lau D, Canobbio M, de Goede A, Singh I, et al. Automated compounding technology and workflow solutions for the preparation of chemotherapy: a systematic review. Eur J Hosp Pharm 2020 Nov 21;27(6):330-336. [doi: 10.1136/ejhpharm-2019-001948] [Medline: 33097615]
2. Baral SD, Poteat T, Strömdahl S, Wirtz AL, Guadamuz TE, Beyrer C. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. Lancet Infect Dis 2013 Mar;13(3):214-222. [doi: 10.1016/S1473-3099(12)70315-8] [Medline: 23260128]
3. Reece K, Lozano M, Roux R, Spivey S. Implementation and evaluation of a gravimetric i.v. workflow software system in an oncology ambulatory care pharmacy. Am J Health Syst Pharm 2016 Feb 01;73(3):165-173. [doi: 10.2146/ajhp150169] [Medline: 26796911]
4. Stamatis D. Failure Mode and Effects Analysis: FMEA From Theory to Execution. Milwaukee, WI: ASQC Quality Press; 1995.
5. ISMP Survey Provides Insights into Pharmacy Sterile Compounding Systems and Practices. 2020 Oct 22. URL: https://ismp.org/resources/ismp-survey-provides-insights-pharmacy-sterile-compounding-systems-and-practices [accessed 2021-10-10]
6. European Medicines Agency. Good Practice Guide on Recording, Coding, Reporting and Assessment of Medication Errors (Report No.: EMA/762563/2014). 2014. URL: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/good-practice-guide-recording-coding-reporting-assessment-medication-errors_en.pdf [accessed 2021-03-15]
7. Ubanyionwu S, Beyer J. Revolutionize medication safety with smart pump-electronic health record interoperability. 2019 Presented at: American Society of Health-System Pharmacists (ASHP) Summer Meetings & Exhibition; June 11, 2019; Boston, MA.
8. Biltoft J, Finneman L. Clinical and financial effects of smart pump-electronic medical record interoperability at a hospital in a regional health system. Am J Health Syst Pharm 2018 Jul 15;75(14):1064-1068. [doi: 10.2146/ajhp161058] [Medline: 29987060]
9. National Pharmacy Association. NPA medication safety update (MSO report) Quarter 4 2019 (England). 2019. URL: https://www.npa.co.uk/news-and-events/news-item/npa-medication-safety-update-mso-report-quarter-4-2019-england/ [accessed 2021-03-01]
10. Waterson J, Al-Jaber R, Kassab T, Al-Jazairi AS. Twelve-Month Review of Infusion Pump Near-Miss Medication and Dose Selection Errors and User-Initiated "Good Save" Corrections: Retrospective Study. JMIR Hum Factors 2020 Aug 11;7(3):e20364 [FREE Full text] [doi: 10.2196/20364] [Medline: 32667895]
11. Independent Double Checks: Worth the Effort if Used Judiciously and Properly. 2019 Jun 6. URL: https://www.ismp.org/resources/independent-double-checks-worth-effort-if-used-judiciously-and-properly [accessed 2021-03-01]
12. Jha AK, Kuperman GJ, Teich JM, Leape L, Shea B, Rittenberg E, et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. J Am Med Inform Assoc 1998;5(3):305-314 [FREE Full text] [doi: 10.1136/jamia.1998.0050305] [Medline: 9609500]
13. Ceriani Cernadas JM, Bogado L, Espínola Rolón F, Galletti MF. Voluntary and anonymous reporting of medication errors in patients admitted to the Department of Pediatrics. Arch Argent Pediatr 2019 Dec 01;117(6):e592-e597 [FREE Full text] [doi: 10.5546/aap.2019.eng.e592] [Medline: 31758886]
14. Briggs B. Point of care can be anywhere. Health Data Manag 2003 Dec;11(12):22-6, 28, 30 passim. [Medline: 14682252]
15. Terkola R, Czejka M, Bérubé J. Evaluation of real-time data obtained from gravimetric preparation of antineoplastic agents shows medication errors with possible critical therapeutic impact: Results of a large-scale, multicentre, multinational, retrospective study. J Clin Pharm Ther 2017 Aug 21;42(4):446-453. [doi: 10.1111/jcpt.12529] [Medline: 28429389]
16. South DA, Skelley JW, Dang M, Woolley T. Near-miss transcription errors: a comparison of reporting rates between a novel error-reporting mechanism and a current formal reporting system. Hosp Pharm 2015 Feb;50(2):118-124 [FREE Full text] [doi: 10.1310/hpj5002-118] [Medline: 25717207]
17. Bekes C. The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. Yearbook of Critical Care Medicine 2007 Jan;2007:319-321. [doi: 10.1016/s0734-3299(08)70434-9]

Abbreviations

AUC: area under the curve
BCMA: barcode medication administration
CPOE: computerized provider order entry
FMEA: failure mode effect analysis
FTE: full-time employee
NMTE: near-miss transcription error

Edited by A Kushniruk; submitted 29.03.21; peer-reviewed by R Terkola, YM Huang; comments to author 13.04.21; revised version received 07.06.21; accepted 25.08.21; published 02.11.21.

Please cite as:
Meren ÜH, Waterson J
Evaluating An Automated Compounding Workflow Software for Safety and Efficiency: Implementation Study
JMIR Hum Factors 2021;8(4):e29180
URL: https://humanfactors.jmir.org/2021/4/e29180
doi:10.2196/29180
PMID:34456182

Original Paper

Using Postmarket Surveillance to Assess Safety-Related Events in a Digital Rehabilitation App (Kaia App): Observational Study

Deeptee Jain[1], MD; Kevin Norman[2], PhD; Zachary Werner[2]; Bar Makovoz[2]; Turner Baker[2]; Stephan Huber[3], MD

[1]Department of Orthopaedic Surgery, Washington University in St. Louis, St. Louis, MO, United States

[2]Neoteric Consulting, New York, NY, United States

[3]Kaia Health GmbH, Munich, Germany

Corresponding Author:
Deeptee Jain, MD
Department of Orthopaedic Surgery
Washington University in St. Louis
660 South Euclid Avenue
St. Louis, MO, 63110
United States
Phone: 1 314 747 4950
Email: deeptee.jain@gmail.com

Abstract

Background: Low back pain (LBP) affects nearly 4 out of 5 individuals during their lifetime and is the leading cause of disability globally. Digital therapeutics are emerging as effective treatment options for individuals experiencing LBP. Despite the growth of evidence demonstrating the benefits of these therapeutics in reducing LBP and improving functional outcomes, little data has been systematically collected on their safety profiles.

Objective: This study aims to evaluate the safety profile of a multidisciplinary digital therapeutic for LBP, the Kaia App, by performing a comprehensive assessment of reported adverse events (AEs) by users as captured by a standardized process for postmarket surveillance.

Methods: All users of a multidisciplinary digital app that includes physiotherapy, mindfulness techniques, and education for LBP (Kaia App) from 2018 to 2019 were included. Relevant messages sent by users via the app were collected according to a standard operating procedure regulating postmarket surveillance of the device. These messages were then analyzed to determine if they described an adverse event (AE). Messages describing an AE were then categorized based on the type of AE, its seriousness, and its relatedness to the app, and they were described by numerical counts. User demographics, including age and gender, and data on app use were collected and evaluated to determine if they were risk factors for increased AE reporting.

Results: Of the 138,337 active users of the Kaia App, 125 (0.09%) reported at least one AE. Users reported 0.00014 AEs per active day on the app. The most common nonserious AE reported was increased pain. Other nonserious AEs reported included muscle issues, unpleasant sensations, headache, dizziness, and sleep disturbances. One serious AE, a surgery, was reported. Details of the event and its connection to the intervention were not obtainable, as the user did not provide more information when asked to do so; therefore, it was considered to be possibly related to the intervention. There was no relationship between gender and AE reporting (P>.99). Users aged 25 to 34 years had reduced odds (odds ratio [OR] 0.31, 95% CI 0.08-0.95; P=.03) of reporting AEs, while users aged 55 to 65 years (OR 2.53, 95% CI 1.36-4.84, P=.002) and ≥75 years (OR 4.36, 95% CI 1.07-13.26; P=.02) had increased odds. AEs were most frequently reported by users who had 0 to 99 active days on the app, and less frequently reported by users with more active days on the app.

Conclusions: This study on the Kaia App provides the first comprehensive assessment of reported AEs associated with real-world use of digital therapeutics for lower back pain. The overall rate of reported AEs was very low, but significant reporting bias is likely to be present. The AEs reported were generally consistent with those described for in-person therapies for LBP.

(JMIR Hum Factors 2021;8(4):e25453) doi:10.2196/25453

KEYWORDS

lower back pain; digital therapeutics; adverse event; pain; safety; digital health; multidisciplinary pain treatment

Introduction

Low back pain (LBP) is the leading cause of long-term pain and physical disability in developed countries [1-3]. Nearly 80% of individuals are affected by LBP during their lifetime [4,5]. LBP imposes a major socioeconomic burden on both individuals and industry [6]. In the United States, lost productivity due to LBP, including an estimated 264 million work days lost annually [7,8], contributes to a total economic burden of LBP that exceeds US $100 billion [9,10].

Evidence-based clinical guidelines recommend nonpharmacological approaches, including exercise and mindfulness-based stress reduction care, for individuals experiencing lower back pain [11]. Multidisciplinary pain treatment programs that supplement physiotherapy with mindfulness, exercises, and educational materials are more efficacious at alleviating long-term LBP than physical therapy alone [12-15]. Traditional in-person treatments, however, have a few limitations. They are often costly, which may limit access to those with lower financial means. Furthermore, physiotherapy programs rely on continuous care between appointments and performing exercises independently at home; this reduces adherence, thereby limiting effective treatment [16,17].

Novel interventions, including digital platforms, are becoming increasingly popular to support medical treatment while addressing the limitations of standard in-person treatment options. Digital therapeutics are products that aim to leverage digital, software, or internet-based health technologies to deliver to prevent, manage, or treat medical disorders [18]. Digital therapeutics provide conventional evidence-based interventions on a highly accessible digital platform and in a continuous manner [18]. Digital approaches for LBP are becoming increasingly popular as a means to use the evidence-based, standard of care physical therapy and mindfulness techniques recommended by physicians while increasing accessibility, maintaining program adherence, and reducing costs for users. Multiple digital therapeutic interventions for LBP have been developed, and previous randomized controlled trials (RCTs) have shown that they are effective for reducing pain and disability indices [19-22] and improving adherence to an exercise program [23].

However, few data are available on the safety of these programs. Despite the ease at which digital therapeutics can allow the streamlined collection and recording of safety data from users, some studies fail to report on adverse events (AEs) [19,20,23], while others do not clearly define the methodologies used for reporting [21,22,24]. Additionally, the small sample sizes of the RCTs may have limited the studies from capturing AEs that occur less frequently. AE reporting is critical to identify potential risks associated with the intervention.

The objective of this study was to evaluate the AEs captured with a systematic vigilance process during real-world use of a specific digital therapeutic for LBP, the Kaia App. It was hypothesized that users of the digital therapeutic for LBP would report similar AEs to those in comparable nondigital programs.

Methods

Study Design

This study examined the adverse event reporting of all users of a multidisciplinary digital app (the Kaia App) that includes physiotherapy, mindfulness techniques, and education for LBP, from 2018-2019. Relevant messages sent by users via the app were collected according to a standard operating procedure (SOP) regulating postmarket surveillance of the device. These messages were then analyzed to determine if they described an AE. Messages describing an AE were then categorized based on the type of AE, seriousness, and relatedness to the app, and they were described by numerical counts. User demographics, including age and gender, and app use data were collected and evaluated to determine if they were risk factors for increased AE reporting.

Participants

This retrospective case series included users who were active on the Kaia App from January 2018 to December 2019. Participants had self-reported low back pain. Onboarding criteria for the program have been previously described [25]. The study population in this study consisted of all international users whose interactions were traceable with Kaia's ticketing system and who were active on the app in 2018 or 2019. Due to data privacy laws, users were given the option to opt in to the use of their personal demographic (age and gender) and app use data, such as active days using the app during the research study. Active days were defined as the number of days in which the users interacted with the app. The rate of AEs per active day of using the app is a metric used to calculate the expected frequency of an AE and to provide a sense of the overall safety profile of the app. This is consistent with the risk management processes for medical devices according to the International Organization for Standardization (ISO 14791). Users were able to withdraw use of the app at any point or to opt out of the collection and storage of personal data.

Ethical Considerations

The study was conducted with a deidentified data set, which did not contain any electronic personal health information. As such, the study was considered institutional review board–exempt by the Institutional Research Board of the Bavarian Regional Medical Council (2020-1198, Bayerische Landesärztekammer).

Kaia App Modules

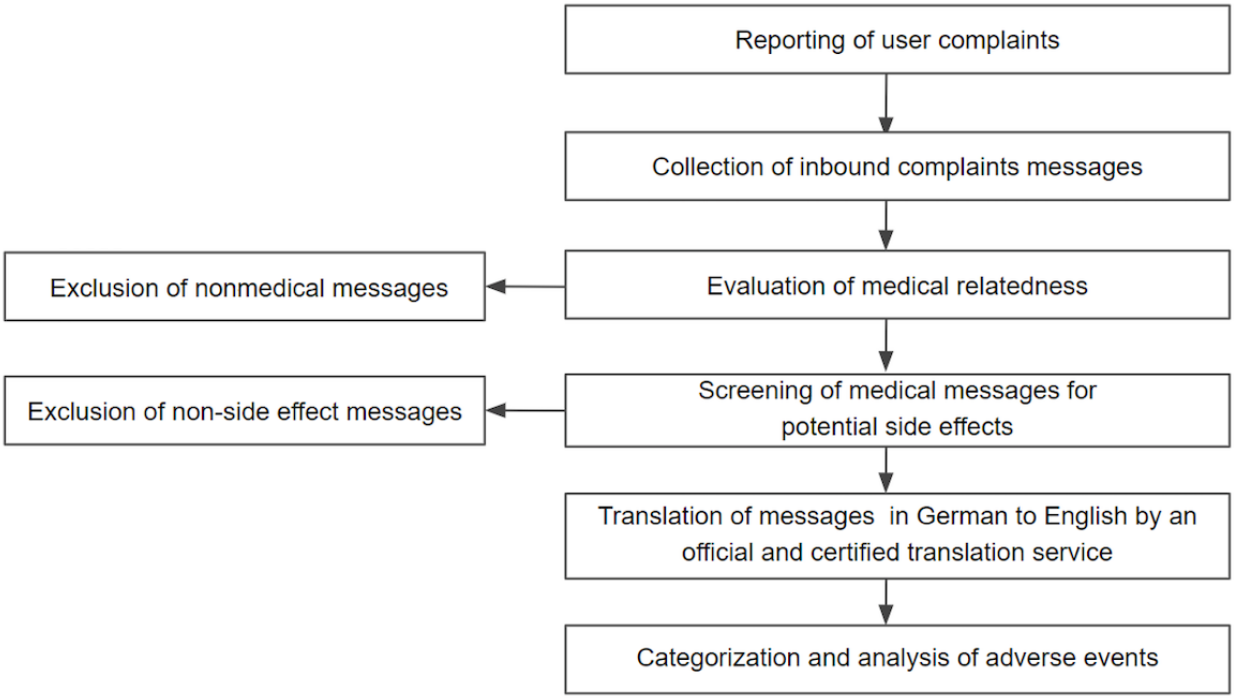
Kaia Health offers a multidisciplinary digital therapeutic solution (Kaia App) for LBP, which has been previously shown to effectively reduce LBP with guided physiotherapy, mindfulness, and educational training [21,26,27]. The Kaia App [25] includes three therapy modules, (1) physiotherapy, (2) mindfulness, and (3) education, with exercises to be performed on a daily basis. The content for each individual user is adapted daily based on the previously completed modules. In this study, users were not obligated to participate in all three modules in a given session. Physiotherapy was limited to up to 5 exercises. The database of 145 exercises was subcategorized into 5 classes based on the targeted body location for that exercise. The exercises

recommended were dependent on where the user indicated the most pain. Recommended exercises were adjusted based on ongoing user feedback.

Reporting of AEs

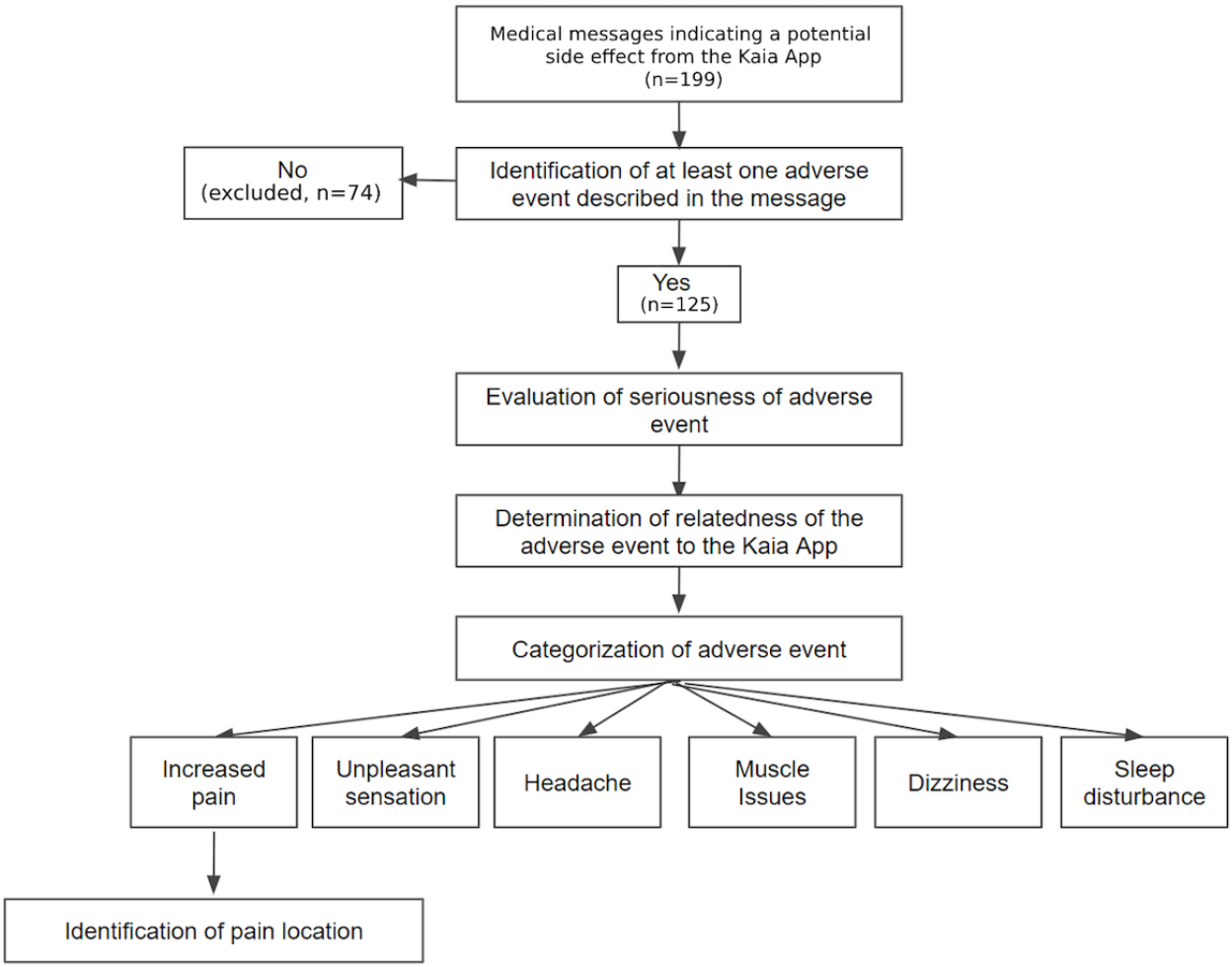
Users regularly corresponded with a personal coach or customer support through the app. Users self-reported AEs to their coach or customer support staff, and the messages were analyzed retrospectively after the users stopped using the app. Users were not specifically prompted to report AEs. All messages indicating potential complaints were tracked in the ticketing system according to an SOP regulating postmarket surveillance of the device (Figure 1). A *complaint* was defined as any written, electronic, or verbal communication that alleged deficiencies related to the identity, quality, durability, reliability, safety, effectiveness, or performance of the app. All messages were screened for medical relatedness and potential side effects by customer support, and if they contained any suggestion of an adverse event, they were forwarded to the Kaia Health medical and quality management team. The process was regulated by the SOP of Kaia Health. The customer support team was trained to label all complaints as either a medical complaint or technical issue. The workflow followed a Kaia internal SOP that includes didactic and supervised learning models. All flagged medical complaints were reviewed by at least one trained, board-certified MD in the field of musculoskeletal pain and a regulatory quality management representative who confirmed each complaint as medically relevant. Any messages written in German were translated to English by a certified translation service, Medax Translation Services (Olching, Germany).

Figure 1. Procedure for the collection and analysis of Kaia App user complaints.



A total of 199 medically related messages indicating potential side effects were identified. These messages were then assessed to determine if they described an AE, defined as any untoward medical event. The seriousness, category, and relatedness of the AE to the app was then evaluated (Figure 2), as described below. Two researchers categorized each of the messages independently. Each researcher was blinded to the other's responses. Discrepancies were decided by a third independent member.

Figure 2. Study process of the medically related message assessment procedure.



Adverse Event Seriousness Assessment

The following definitions were used to classify the seriousness of an AE [28]. A serious AE was defined as any untoward medical occurrence that resulted in one of the following outcomes: death, illness/injury requiring hospitalization, events deemed life-threatening, or significant disability. All other AEs were considered nonserious, whether or not they were considered to be related to the intervention. Messages were not considered AEs if they only contained updates on progress, inquiries, or advice on app use.

Adverse Event Categorization

Previous literature identifying categories of AEs related to exercise and pain management was used to create the classification of AEs [29-31]. The following categories were determined: increased pain, muscle issues, headache, dizziness, unpleasant sensation, and sleep disturbances. Increased pain included any indications of increased pain compared to the user's normal pain level. Muscle issues included muscle-specific discomfort, such as the reporting of muscle cramps, soreness, stiffness, or tightness. Headache included pain in any region of the head. Unpleasant sensation included any reporting of abnormal, uncomfortable sensations, including a feeling of “pins and needles” or unpleasant back cracking noises. Sleep disturbances included disrupted patterns of sleep, including waking up in the middle of the night or difficulty falling asleep.

AE Relatedness Assessment

To assess the relatedness of reported adverse events, we followed the best practices for AE reporting to the US Food and Drug Administration by registries of postmarket products and applied these principles to the Kaia App digital therapeutic [32,33]. There is no standard nomenclature for describing this relationship, as previous studies have used a variety of terms, such as certainly, definitely, probably, possibly, or likely related or not related [33]. In this study, an AE was considered related to the app intervention if the AE was (1) a known response to similar interventions (ie, biological plausibility) and (2) temporally linked to the intervention. AEs were categorized as definitely or possibly related to app use. AEs were considered definitely related if there was a reasonable, temporal relationship between the AE and the intervention, the AE was consistent with a known or expected response pattern to the intervention, and the AE could not be reasonably explained by the known characteristics of the user’s clinical state. AEs were considered possibly related if the AE followed a reasonable temporal sequence from administration of the study intervention and followed a known or expected response pattern to the intervention, but that could readily have been produced by a number of other factors. If AEs had vague or ambiguous

temporal relationships with app use or might reasonably have been a result of a pre-existing condition described in the message, the AE was identified as possibly related.

Pain Location

The majority of user messages indicating an AE of increased pain specifically identified the location of the pain on the body. AEs indicating increased pain were subcategorized based on location, including pain in the back (including indication of upper and lower back pain and sacroiliac joint pain), neck, shoulder, leg or knee, or other regions, including sciatica, hip pain, or arm pain. If the message did not mention the location of the pain, it was considered nonspecified.

Statistical Analyses

All statistical tests were performed using RStudio, version 3.5.3 (R Foundation for Statistical Computing). Demographic variables of total app users and users who reported an AE were described by frequency and as distribution (%) within the group. For age variables, odds relative to the age range of 45 to 54 years and 95% confidence intervals were calculated [34]. The relationships between variables (gender, age, and active days) and AEs were analyzed using the Fisher exact test. A 2-sided *P* value <.05 was used to determine statistical significance.

Results

Overview of Adverse Event Reporting

A summary of AE reporting from users of the Kaia App for back pain is provided in Table 1. A total of 138,337 users were included. Of the 199 medical-related messages sent by users, 125 reported an AE. These 125 users (0.09% of the total population of 138,337) reported a total of 142 AEs. Among all users in the total population, the app was used for 1,004,430 active days. The rate of AEs was 0.00014 per active day.

Table 1. Overview of adverse event reporting on the Kaia App.

Characteristic	Value
Total users on app, N	138,337
Total users reporting an adverse event, n (%)	125 (0.09)
Total adverse events reported, n	142
Total active days using the app, n	1,004,430
Rate of reported adverse events per active day	0.00014

Demographics

The genders of the all users and the users reporting an AE are displayed in Table 2. Demographic data were available for 74 of the 125 users who reported an AE. No relationship between gender and the reporting of AEs was found (Fisher exact test, *P*>.99).

The ages of all users and users reporting an AE are displayed in Table 3. An odds ratio and 95% confidence interval for AEs was calculated for each age group relative to the age group with the largest number of users (ages 45-54 years). Individuals aged 25-34 years had reduced odds (*P*=.03) of reporting AEs, while those aged 55-65 years (*P*=.002) and ≥75 years (*P*=.02) had increased odds (Fisher exact test).

Table 2. Gender demographics of the app users (N=138,337). Demographic data were available for 74 of the 125 users who reported an adverse event.

Gender	Value, n (%)	
	All users	Users reporting an adverse event
Female	76,906 (55.6)	42 (56.8)
Male	57,152 (41.3)	31 (41.9)
Unspecified	4279 (3.1)	1 (1.4)

Table 3. Relationship between age and adverse events.

Age (years)	Values				
	All users (N=138,337), n (%)	Users reporting an adverse event (n=74), n (%)	Odds ratio	95% CI	*P* value
<25	9369 (6.8)	1 (1.4)	0.21	0.01-1.35	.15
25-34	25,531 (18.5)	4 (5.4)	0.31	0.08-0.95	.03[a]
35-44	34,826 (25.2)	18 (24.3)	1.20	0.61-2.39	.63
45-54	35,847 (25.9)	15 (20.3)	Reference	Reference	Reference
55-64	22,824 (16.5)	26 (35.1)	2.53	1.36-4.84	.002[a]
65-75	8089 (5.8)	8 (10.8)	1.97	0.74-4.77	.13
>75	1829 (1.3)	2 (2.7)	4.36	1.07-13.26	.02[a]

[a]*P*<.05.

Categories of Adverse Events Reported and Relationship with App Use

The specific categories of reported AEs are shown in Table 4. Most of the AEs were nonserious, including increased pain, muscle issues, unpleasant sensations, headache, dizziness, and sleep disturbances. All nonserious AEs were determined to be either possibly or definitely related to the digital intervention. One user reported a serious AE, a surgery that occurred during the time period when the individual was using the intervention. Given that we do not have additional information beyond the user messages, we do not know what kind of surgery was performed. This serious AE was rated as possibly related to use of the digital intervention. Users were using the Kaia App as a therapeutic for back pain; therefore, it is uncertain that the injury resulting in surgery was a pre-existing cause of the user's original back pain or a new symptom.

The anatomical location in which users reported increased pain was then categorized, as shown in Table 5. Due to the self-reporting nature of the AE reporting, many of the users experiencing increased pain did not report the specific location of the increased pain. Of the users who did report a location, back pain was the most common location reported. Users also experienced increased pain in the lower extremities (leg or knee), shoulder, neck, or other body parts including the hip and arms.

Finally, the relationship between the number of active days on the app and the frequency of reported AEs is examined in Table 6. The average number of active days per app user of the total cohort was 7.26 days. AEs were most frequently reported by users who had 0 to 99 active days on the app and less frequently reported by users with more active days on the app.

Table 4. Adverse events per category type.

Category of adverse event	Frequency (n=142), n (%)
Increased pain	83 (58.4)
Muscle issues	25 (17.6)
Unpleasant sensation	19 (13.4)
Headache	7 (4.9)
Dizziness	4 (2.8)
Sleep disturbance	3 (2.1)
Surgery	1 (0.7)

Table 5. Total adverse events reported per location of increased pain.

Location of increased pain	Frequency (n=83), n (%)
Back	25 (30.1)
Leg or knee	11 (13.2)
Shoulder	11 (13.2)
Neck	8 (9.6)
Other	8 (9.6)
Not specified	27 (32.5)

Table 6. Total adverse events reported per active days using the Kaia App. App use data were available for 84 of the 125 users who reported an adverse event.

Active days on Kaia App	Adverse events, n (%)
0-99	51 (60.7)
100-199	18 (21.4)
200-299	6 (7.1)
300-399	6 (7.1)
400-499	2 (2.4)
500-599	1 (1.2)

Discussion

Principal Findings

This study provides the first comprehensive assessment of reported AEs associated with real-world use of a digital therapeutic for LBP (the Kaia App). In this retrospective case series, only 0.9% of users reported an AE. AEs were mostly nonserious and included increased pain, muscle issues, dizziness, headaches, and sleep disturbances. The back was the most common location of increased pain reported by users of the app. One serious adverse event, a surgery, was reported; it was determined to be possibly related to the digital intervention, as it could not be determined whether the cause of the surgery was due to the intervention or the underlying condition.

There was no relationship between gender and the reporting of adverse events. Younger users had reduced odds of reporting AEs, while older users had increased odds. On average, users only reported 0.00014 adverse events per active day using the app.

Comparison With Prior Work

Randomized controlled trials evaluating the use of digital therapeutics for lower back pain have included limited AE reporting [19-24] (Table S1 in Multimedia Appendix 1). The table includes randomized controlled trials that evaluated the use of digital therapeutics for lower back pain and included an analysis of their adverse event reporting. These trials were identified through using comprehensive search terms across the MEDLINE, Embase and Web of Science databases to collect all trials that assessed the use of telehealth interventions available to at-home patients. We found that most of these prior studies did not provide detailed reporting of adverse events; therefore, it is challenging to directly compare the safety of the LBP digital therapeutic in this study to that of other digital-based programs for LBP management.

The AEs reported in this study are comparable to those reported for nondigital forms of the three therapy modules included in the app, including (1) physiotherapy, (2) mindfulness and relaxation exercises, and (3) education for LBP.

All of the AEs in this study were consistent with previously reported AEs related to live exercise therapy. Exercise intervention, while considered safe overall, has been shown to increase the risk of experiencing nonserious AEs in individuals with LBP, but not of serious AEs [35]. Participants who perform either back-focused physical therapy exercises or yoga for LBP [31,36-39] report more AEs than control participants who perform less strenuous nonexercises [40,41]. Most previously reported AEs associated with exercise therapy are musculoskeletal in nature, including increased pain [35] and muscle soreness [42], as well as other nonserious AEs [43] such as headache and dizziness. Previously reported AEs related to yoga include joint pain, increased back pain, sciatica or leg pain, neck pain, abdominal pain, and dizziness [31]. Of note, the rate (0.09%) of reported AEs in this study with a digital app was much lower than what has been reported in prior studies of live exercise therapy for LBP, such as physical therapy (7%-11%) [44] or yoga (7.1%-7.6%) [31,45], although this finding may be limited by the self-reported nature of the AEs collected in this study.

LBP is the second most common reason to visit a primary care physician; it is self-identified, and it is the chief concern upon presentation [46-48]. Thus, the fact that users in this study self-identify as having low back pain makes this study widely generalizable to a broad population.

The pain-related AEs reported in this study have also been reported in prior literature examining mindfulness exercises, but the risk is low. A previous study reported that 10% of individuals with chronic LBP experienced an AE during cognitive behavioral therapy, which was mostly attributed to increased pain from progressive muscle relaxation exercises [49]. However, progressive muscle relaxation techniques have been demonstrated to result in no AEs in individuals with chronic neck pain [50], suggesting that the location of pain before starting the module may influence AE reporting. No study that specifically examined the relationship between breathing exercises and AEs was found.

Finally, the increased number of LBP AEs seen in this study is also consistent with prior literature examining education material related to LBP; however, the risk is low. Literature searches reveal that the existing AE reporting for these interventions is limited, as they are generally considered safe. Individuals given self-care books and newsletters that recommend nonstrenuous stretching routines report very low rates (1.6%) of adverse events, including increased back pain [31]. In another study of participants with LBP assigned to a self-care book treatment, 2.2% participants reported an AE of increased back pain [38].

Next, this study sought to identify risk factors, including age and gender, that may be associated with increased likelihood of reporting an AE. We found that increasing age was a risk factor for reporting an AE. Although moderate to intense

exercise has been shown to be safe overall in a healthy population of older people [51-53], older individuals are indeed at increased risk for injury from falls during physical activity [54]. We did not identify a relationship with gender and AE reporting in this study, and to our knowledge, this is the first study to examine gender differences in AE reporting on physical therapy. Although our study focused on demographic risk factors for reporting an AE while using the Kaia App for LBP, future studies should examine additional aspects of back pain that could be risk factors, such as the intensity, duration, and history of LBP [55].

Limitations

The major limitation of this study is that it was retrospective and relied on self-reporting of possible AEs. Users self-reported AEs to their coach or customer support staff, and the messages were analyzed retrospectively after users stopped using the app. Users were not specifically prompted to report AEs. Overall, this likely resulted in underreporting, and it may explain the low rate of AEs in this study compared to that in prior studies examining live physiotherapy [31]. In particular, the ability to self-report serious AEs is inherently flawed, and more accurate results on the incidence of those events can be more optimally obtained from a prospective study cohort, where planned follow-ups will accurately collect those events. Serious AEs, such as death, cannot be reported by the user, as they would be unable to use the app to report any such event. Although this underreporting would be a serious concern in the tracking of high-risk interventions, by nature, the described intervention is extremely unlikely to cause death or serious AE. Another limitation is that users submitted open-ended messages of variable length and detail to their coach; thus, categorization of AEs was subjective. Although users were not specifically prompted to state the temporal relationship between the AE and app use, many messages did indicate this relationship. To mitigate these issues, two independent researchers classified the medical complaints separately using strict definitions for AE categorization and relatedness, and a third researcher made the final decision on the classification in the event of a discrepancy between the two initial reviewers. To better understand AEs in digital therapeutics, we recommend that a prospective study design be implemented, that users be prompted frequently to report AEs, and that reporting of AEs trigger follow-up questions regarding details.

Finally, due to data privacy laws, users were given the option to opt in to the use of their personal demographic (age and gender) and app use data by the manufacturer; thus, data were missing for some patients who reported adverse events. This may have impacted the analysis of demographics and app use on AE reporting.

Conclusions

This study serves to emphasize the importance of examining AEs in digital therapeutics for LBP, as these therapeutics are becoming an increasingly popular treatment modality. Future research on digital LBP rehabilitation tools should focus on prospective assessment of AEs using the streamlined nature of data collection in digital interventions to gather safety data from users to identify potential risk factors for negative health consequences.

Acknowledgments

The authors would like to thank Medax Translation Services for the translation of German language messages into English.

Conflicts of Interest

This study was funded by Kaia Health. DJ and TB receive compensation for advisory services. KN, ZW, and BM receive compensation for advisory services through Neoteric Consulting Group. SH is an employee of Kaia Health and receives stock options and salary.

Multimedia Appendix 1
Review of the adverse event reporting of randomized controlled trials on digital therapeutics for lower back pain management.
[DOCX File , 15 KB - humanfactors_v8i4e25453_app1.docx]

References

1. Woolf AD, Erwin J, March L. The need to address the burden of musculoskeletal conditions. Best Pract Res Clin Rheumatol 2012 Apr;26(2):183-224. [doi: 10.1016/j.berh.2012.03.005] [Medline: 22794094]
2. Health in the European Union. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Health_in_the_European_Union_–_facts_and_figures [accessed 2020-12-20]
3. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. Lancet 2018 Jun;391(10137):2356-2367. [doi: 10.1016/s0140-6736(18)30480-x]
4. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 2018 Nov;392(10159):1789-1858. [doi: 10.1016/S0140-6736(18)32279-7]
5. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet 2017 Feb;389(10070):736-747. [doi: 10.1016/s0140-6736(16)30970-9]

6. Taimela S, Kujala UM, Salminen JJ, Viljanen T. The prevalence of low back pain among children and adolescents. A nationwide, cohort-based questionnaire survey in Finland. Spine (Phila Pa 1976) 1997 May 15;22(10):1132-1136. [doi: 10.1097/00007632-199705150-00013] [Medline: 9160472]
7. Guo HR, Tanaka S, Halperin WE, Cameron LL. Back pain prevalence in US industry and estimates of lost workdays. Am J Public Health 1999 Jul;89(7):1029-1035. [doi: 10.2105/ajph.89.7.1029] [Medline: 10394311]
8. The Burden of Musculoskeletal Diseases in the United States: Prevalence, Societal and Economic Cost. URL: https://www.boneandjointburden.org/docs/BMUS%20Impact%20of%20MSK%20on%20Americans%20booklet_4th%20Edition%20%282018%29.pdf [accessed 2020-12-01]
9. Katz J. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. J Bone Joint Surg Am 2006 Apr;88 Suppl 2:21-24. [doi: 10.2106/JBJS.E.01273] [Medline: 16595438]
10. Rubin DI. Epidemiology and risk factors for spine pain. Neurol Clin 2007 May;25(2):353-371. [doi: 10.1016/j.ncl.2007.01.004] [Medline: 17445733]
11. Qaseem A, Wilt TJ, McLean RM, Forciea MA. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. Ann Intern Med 2017 Feb 14;166(7):514. [doi: 10.7326/m16-2367]
12. Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. Pain 1992 May;49(2):221-230. [doi: 10.1016/0304-3959(92)90145-2] [Medline: 1535122]
13. Marin T, Van Eerd D, Irvin E, Couban R, Koes BW, Malmivaara A, et al. Multidisciplinary biopsychosocial rehabilitation for subacute low back pain. Cochrane Database Syst Rev 2017 Jun 28;6:CD002193 [FREE Full text] [doi: 10.1002/14651858.CD002193.pub2] [Medline: 28656659]
14. Patrick LE, Altmaier EM, Found EM. Long-term outcomes in multidisciplinary treatment of chronic low back pain: results of a 13-year follow-up. Spine (Phila Pa 1976) 2004 Apr 15;29(8):850-855. [doi: 10.1097/00007632-200404150-00006] [Medline: 15082983]
15. Cassidy E, Atherton RJ, Robertson N, Walsh DA, Gillett R. Mindfulness, functioning and catastrophizing after multidisciplinary pain management for chronic low back pain. Pain 2012 Mar;153(3):644-650. [doi: 10.1016/j.pain.2011.11.027] [Medline: 22240149]
16. Palazzo C, Klinger E, Dorner V, Kadri A, Thierry O, Boumenir Y, et al. Barriers to home-based exercise program adherence with chronic low back pain: patient expectations regarding new technologies. Ann Phys Rehabil Med 2016 Apr;59(2):107-113 [FREE Full text] [doi: 10.1016/j.rehab.2016.01.009] [Medline: 27050664]
17. Jack K, McLean SM, Moffett JK, Gardiner E. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228 [FREE Full text] [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
18. Dang A, Arora D, Rane P. Role of digital therapeutics and the changing future of healthcare. J Family Med Prim Care 2020 May;9(5):2207-2213 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_105_20] [Medline: 32754475]
19. Chhabra HS, Sharma S, Verma S. Smartphone app in self-management of chronic low back pain: a randomized controlled trial. Eur Spine J 2018 Nov 15;27(11):2862-2874. [doi: 10.1007/s00586-018-5788-5] [Medline: 30324496]
20. Shebib R, Bailey JF, Smittenaar P, Perez DA, Mecklenburg G, Hunter S. Randomized controlled trial of a 12-week digital care program in improving low back pain. NPJ Digit Med 2019 Jan 7;2(1):1 [FREE Full text] [doi: 10.1038/s41746-018-0076-7] [Medline: 31304351]
21. Toelle TR, Utpadel-Fischler DA, Haas K, Priebe JA. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. NPJ Digit Med 2019 May 3;2(1):34 [FREE Full text] [doi: 10.1038/s41746-019-0109-x] [Medline: 31304380]
22. Zadro J, Shirley D, Simic M, Mousavi SJ, Ceprnja D, Maka K, et al. Video-game-based exercises for older people with chronic low back pain: a randomized controlledtable trial (GAMEBACK). Phys Ther 2019 Jan 01;99(1):14-27. [doi: 10.1093/ptj/pzy112] [Medline: 30247715]
23. Geraghty AWA, Stanford R, Stuart B, Little P, Roberts LC, Foster NE, et al. Using an internet intervention to support self-management of low back pain in primary care: findings from a randomised controlled feasibility trial (SupportBack). BMJ Open 2018 Mar 09;8(3):e016768 [FREE Full text] [doi: 10.1136/bmjopen-2017-016768] [Medline: 29525768]
24. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. J Pain 2016 Dec;17(12):1302-1317 [FREE Full text] [doi: 10.1016/j.jpain.2016.08.011] [Medline: 27616607]
25. Huber S, Priebe JA, Baumann K, Plidschun A, Schiessl C, Tölle TR. Treatment of low back pain with a digital multidisciplinary pain treatment app: short-term results. JMIR Rehabil Assist Technol 2017 Dec 04;4(2):e11 [FREE Full text] [doi: 10.2196/rehab.9032] [Medline: 29203460]
26. Priebe JA, Haas KK, Moreno Sanchez LF, Schoefmann K, Utpadel-Fischler DA, Stockert P, et al. Digital treatment of back pain versus standard of care: the cluster-randomized controlled trial, Rise-uP. J Pain Res 2020 Jul;13:1823-1838. [doi: 10.2147/jpr.s260761]
27. Priebe JA, Utpadel-Fischler D, Toelle T. Less pain, better sleep? The effect of a multidisciplinary back pain app on sleep quality in individuals suffering from back pain – a secondary analysis of app user data. J Pain Res 2020 May;Volume 13:1121-1128. [doi: 10.2147/jpr.s232792]

28. CFR - Code of Federal Regulations Title 21. US Food and Drug Administration. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm [accessed 2020-12-01]
29. Krein SL, Abdul-Wahab Y, Kadri R, Richardson CR. Adverse events experienced by participants in a back pain walking intervention: a descriptive study. Chronic Illn 2016 Mar 19;12(1):71-80. [doi: 10.1177/1742395315601416] [Medline: 26289360]
30. Dworkin RH, Jensen MP, Gould E, Jones BA, Xiang Q, Galer BS, et al. Treatment satisfaction in osteoarthritis and chronic low back pain: the role of pain, physical and emotional functioning, sleep, and adverse events. J Pain 2011 Apr;12(4):416-424. [doi: 10.1016/j.jpain.2010.07.009] [Medline: 20869325]
31. Saper RB, Lemaster C, Delitto A, Sherman KJ, Herman PM, Sadikova E, et al. Yoga, physical therapy, or education for chronic low back pain. Ann Intern Med 2017 Jun 20;167(2):85. [doi: 10.7326/m16-2579]
32. Gliklich RE, Leavy MB, Dreyer NA, editors. Registries for Evaluating Patient Outcomes: A User's Guide. Rockville, MD: Agency for Healthcare Research and Quality; Sep 01, 2020.
33. Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment. US Food and Drug Administration. 2005. URL: https://www.fda.gov/files/drugs/published/Good-Pharmacovigilance-Practices-and-Pharmacoepidemiologic-Assessment-March-2005.pdf [accessed 2020-12-01]
34. Tenny S, Hoffman MR. Odds ratio. In: StatPearls. Treasure Island, FL: StatPearls Publishing; Jan 01, 2020.
35. Niemeijer A, Lund H, Stafne SN, Ipsen T, Goldschmidt CL, Jørgensen CT, et al. Adverse events of exercise therapy in randomised controlled trials: a systematic review and meta-analysis. Br J Sports Med 2020 Sep 28;54(18):1073-1080. [doi: 10.1136/bjsports-2018-100461] [Medline: 31563884]
36. Nambi G, Inbasekaran D, Khuman R, Devi S, Shanmugananth, Jagannathan K. Changes in pain intensity and health related quality of life with Iyengar yoga in nonspecific chronic low back pain: A randomized controlled study. Int J Yoga 2014 Jan;7(1):48-53 [FREE Full text] [doi: 10.4103/0973-6131.123481] [Medline: 25035607]
37. Tekur P, Singphow C, Nagendra HR, Raghuram N. Effect of short-term intensive yoga program on pain, functional disability and spinal flexibility in chronic low back pain: a randomized control study. J Altern Complement Med 2008 Jul;14(6):637-644. [doi: 10.1089/acm.2007.0815] [Medline: 18673078]
38. Sherman KJ, Cherkin DC, Wellman RD, Cook AJ, Hawkes RJ, Delaney K, et al. A randomized trial comparing yoga, stretching, and a self-care book for chronic low back pain. Arch Intern Med 2011 Dec 12;171(22):2019-2026 [FREE Full text] [doi: 10.1001/archinternmed.2011.524] [Medline: 22025101]
39. Wieland LS, Santesso N. A summary of a Cochrane Review: yoga treatment for chronic non-specific low back pain. Eur J Integr Med 2017 Apr;11:39-40 [FREE Full text] [doi: 10.1016/j.eujim.2017.03.003] [Medline: 29057019]
40. Saper R, Sherman KJ, Cullum-Dugan D, Davis RB, Phillips RS, Culpepper L. Yoga for chronic low back pain in a predominantly minority population: a pilot randomized controlled trial. Altern Ther Health Med 2009;15(6):18-27 [FREE Full text] [Medline: 19943573]
41. Williams K, Petronis J, Smith D, Goodrich D, Wu J, Ravi N, et al. Effect of Iyengar yoga therapy for chronic low back pain. Pain 2005 May;115(1-2):107-117. [doi: 10.1016/j.pain.2005.02.016] [Medline: 15836974]
42. McKnight PE, Kasle S, Going S, Villanueva I, Cornett M, Farr J, et al. A comparison of strength training, self-management, and the combination for early osteoarthritis of the knee. Arthritis Care Res (Hoboken) 2010 Jan 15;62(1):45-53 [FREE Full text] [doi: 10.1002/acr.20013] [Medline: 20191490]
43. Maiers M, Evans R, Hartvigsen J, Schulz C, Bronfort G. Adverse events among seniors receiving spinal manipulation and exercise in a randomized clinical trial. Man Ther 2015 Apr;20(2):335-341. [doi: 10.1016/j.math.2014.10.003] [Medline: 25454683]
44. Chilibeck PD, Vatanparast H, Cornish SM, Abeysekara S, Charlesworth S. Evidence-based risk assessment and recommendations for physical activity: arthritis, osteoporosis, and low back pain. Appl Physiol Nutr Metab 2011 Jul;36 Suppl 1(S1):S49-S79 [FREE Full text] [doi: 10.1139/h11-037] [Medline: 21800948]
45. Tilbrook HE, Cox H, Hewitt CE, Kang'ombe AR, Chuang LH, Jayakody S, et al. Yoga for chronic low back pain: a randomized trial. Ann Intern Med 2011;155(9):569-578. [doi: 10.1016/j.spinee.2013.02.052]
46. Finley CR, Chan DS, Garrison S, Korownyk C, Kolber MR, Campbell S, et al. What are the most common conditions in primary care? Systematic review. Can Fam Physician 2018 Nov;64(11):832-840 [FREE Full text] [Medline: 30429181]
47. McPhillips-Tangum CA, Cherkin DC, Rhodes LA, Markham C. Reasons for repeated medical visits among patients with chronic back pain. J Gen Intern Med 1998 May;13(5):289-295 [FREE Full text] [doi: 10.1046/j.1525-1497.1998.00093.x] [Medline: 9613883]
48. Cypress BK. Characteristics of physician visits for back symptoms: a national perspective. Am J Public Health 1983 Apr;73(4):389-395. [doi: 10.2105/ajph.73.4.389] [Medline: 6219588]
49. Cherkin DC, Sherman KJ, Balderson BH, Cook AJ, Anderson ML, Hawkes RJ, et al. Effect of mindfulness-based stress reduction vs cognitive behavioral therapy or usual care on back pain and functional limitations in adults with chronic low back pain: a randomized clinical trial. JAMA 2016 Mar 22;315(12):1240-1249 [FREE Full text] [doi: 10.1001/jama.2016.2323] [Medline: 27002445]

50. Lauche R, Materdey S, Cramer H, Haller H, Stange R, Dobos G, et al. Effectiveness of home-based cupping massage compared to progressive muscle relaxation in patients with chronic neck pain—a randomized controlled trial. PLoS One 2013 Jun 7;8(6):e65378 [FREE Full text] [doi: 10.1371/journal.pone.0065378] [Medline: 23762355]
51. Church TS, Earnest CP, Skinner JS, Blair SN. Effects of different doses of physical activity on cardiorespiratory fitness among sedentary, overweight or obese postmenopausal women with elevated blood pressure: a randomized controlled trial. JAMA 2007 May 16;297(19):2081-2091. [doi: 10.1001/jama.297.19.2081] [Medline: 17507344]
52. Dunn AL, Marcus BH, Kampert JB, Garcia ME, Kohl HW, Blair SN. Comparison of lifestyle and structured interventions to increase physical activity and cardiorespiratory fitness: a randomized trial. JAMA 1999 Jan 27;281(4):327-334. [doi: 10.1001/jama.281.4.327] [Medline: 9929085]
53. King AC, Baumann K, O'Sullivan P, Wilcox S, Castro C. Effects of moderate-intensity exercise on physiological, behavioral, and emotional responses to family caregiving: a randomized controlled trial. J Gerontol A Biol Sci Med Sci 2002 Jan 01;57(1):M26-M36. [doi: 10.1093/gerona/57.1.m26] [Medline: 11773209]
54. Karlsson MK, Vonschewelov T, Karlsson C, Cöster M, Rosengen BE. Prevention of falls in the elderly: a review. Scand J Public Health 2013 Jul 03;41(5):442-454. [doi: 10.1177/1403494813483215] [Medline: 23554390]
55. Williams C, Hancock M, Maher C, McAuley J, Lin C, Latimer J. Predicting rapid recovery from acute low back pain based on the intensity, duration and history of pain: a validation study. Eur J Pain 2014 Sep 19;18(8):1182-1189. [doi: 10.1002/j.1532-2149.2014.00467.x] [Medline: 24648103]

Abbreviations

AE: adverse event
LBP: low back pain
OR: odds ratio
RCT: randomized controlled trial
SOP: standard operating procedure

Edited by A Kushniruk; submitted 05.11.20; peer-reviewed by L Sandal, F Abujarad; comments to author 16.12.20; revised version received 11.02.21; accepted 23.07.21; published 09.11.21.

Please cite as:
Jain D, Norman K, Werner Z, Makovoz B, Baker T, Huber S
Using Postmarket Surveillance to Assess Safety-Related Events in a Digital Rehabilitation App (Kaia App): Observational Study
JMIR Hum Factors 2021;8(4):e25453
URL: https://humanfactors.jmir.org/2021/4/e25453
doi:10.2196/25453
PMID:34751664

Review

Including the Reason for Use on Prescriptions Sent to Pharmacists: Scoping Review

Kathryn Mercer[1,2], MI, PhD; Caitlin Carter[2,3], BA, MLIS; Catherine Burns[1,4], PhD, PEng; Ryan Tennant[1,4], BASc, MASc; Lisa Guirguis[5], BSc Pharm, PhD; Kelly Grindrod[3], PharmD, MSc

[1]Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

[2]Library, University of Waterloo, Waterloo, ON, Canada

[3]School of Pharmacy, University of Waterloo, Waterloo, ON, Canada

[4]Advanced Interface Design Lab, Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

[5]Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Kathryn Mercer, MI, PhD
Library
University of Waterloo
200 University Avenue West
DC 1555
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567 ext 42659
Email: kmercer@uwaterloo.ca

Abstract

Background: In North America, although pharmacists are obligated to ensure prescribed medications are appropriate, information about a patient's reason for use is not a required component of a legal prescription. The benefits of prescribers including the reason for use on prescriptions is evident in the current literature. However, it is not standard practice to share this information with pharmacists.

Objective: Our aim was to characterize the research on how including the reason for use on a prescription impacts pharmacists.

Methods: We performed an interdisciplinary scoping review, searching literature in the fields of health care, informatics, and engineering. The following databases were searched between December 2018 and January 2019: PubMed, Institute of Electrical and Electronics Engineers (IEEE), Association for Computing Machinery (ACM), International Pharmaceutical Abstracts (IPA), and EMBASE.

Results: A total of 3912 potentially relevant articles were identified, with 9 papers meeting the inclusion criteria. The studies used different terminology (eg, indication, reason for use) and a wide variety of study methodologies, including prospective and retrospective observational studies, randomized controlled trials, and qualitative interviews and focus groups. The results suggest that including the reason for use on a prescription can help the pharmacist catch more errors, reduce the need to contact prescribers, support patient counseling, impact communication, and improve patient safety. Reasons that may prevent prescribers from adding the reason for use information are concerns about workflow and patient privacy.

Conclusions: More research is needed to understand how the reason for use information should be provided to pharmacists. In the limited literature to date, there is a consensus that the addition of this information to prescriptions benefits patient safety and enables pharmacists to be more effective. Future research should use an implementation science or theory-based approach to improve prescriber buy-in and, consequently, adoption.

(JMIR Hum Factors 2021;8(4):e22325) doi:10.2196/22325

KEYWORDS

patient safety; human factors; patient engagement; multidisciplinary

Introduction

Medications are generally prescribed for conditions and illnesses for 4 reasons: to cure, to prevent, to slow progression, or to manage symptoms. Drugs can also be prescribed to help diagnose or manage the adverse effects caused by another medication or treatment, often referred to as *off-label* use. Sometimes the reason for use is apparent, such as using oral isotretinoin to treat nodular acne. Other times the reason for use is less apparent, such as using a hypertension medication to treat nightmares related to posttraumatic stress disorder [1].

To fill in the gaps when the reason for use information is not accessible, pharmacists must often rely on the patients to provide the reason for use information [2,3]. Yet, the accuracy of patient's self-reported diagnosis varies widely. While the accuracy is quite good with conditions such as diabetes, it is very low for conditions such as rheumatoid arthritis or heart failure [4-6]. People who have difficulty communicating their diagnoses tend to be older, live with more chronic illness, and have a higher risk of death [7]. This puts the onus on the patient to correctly share the physician's prescribing rationale and amplifies the risk for more vulnerable patients.

In the patient safety literature, there appears to be a consensus that it is safer for pharmacists to have access to a prescription's reason for use [8]. While 80% of hospitals in the United States that have adopted some form of an electronic health record allow pharmacists to interact with the system to view laboratory tests and diagnoses, the reason for use is not identified as a core measure included in the electronic health record [7]. ePrescribing has facilitated the accuracy of prescriptions and some discussion with systems; however, many jurisdictions, including Canada, have not yet adopted this technology due to legislative or cost issues. Therefore, while pharmacists may have access to the patient's health information used by the prescriber to determine the reason for use, they must infer the reason without its explicit inclusion on the patient's record. In contrast to community pharmacies where access to national electronic health record data is only available in some countries and regions, communication of the reason for use remains both a desire of pharmacists and a challenge for health care systems [3,8,9].

Most prescriptions today are written electronically [10]. With the potential for timely access by prescribers and pharmacists, digital prescription records could support the communication of a prescription's reason for use along with the right design. Schiff et al [10] tested an indication-based prescribing system that makes it easier for prescribers to share a prescription's reason for use. In their electronic prescribing system, prescribers start with a diagnosis or problem and then select a treatment option from a list of recommendations. The system would additionally provide suggestions based on a patient's health history, but still allow for complete autonomy of a prescribers' selection [10]. However, there still appears to be very little information on how to include a prescription's reason for use to support pharmacist's decision making.

The objective of this interdisciplinary scoping review is to characterize the research on how including the reason for use on a prescription impacts pharmacists. Given that this topic spans multiple disciplines, the first step is to map relevant literature to identify the potential size and scope. Our goals were to describe the research on the design, implementation, and evaluation of the reason for use information for pharmacists, including the types and sources of evidence, and the areas where further research is needed. When literature on a particular topic is scattered through different disciplines, there is a real risk that the research will be siloed and will not reach those who are in a position to translate the research into practice. Thus, we also aimed to provide health care, informatics, and engineering researchers with a cohesive summary of the reason for use studies to date, as it relates to pharmacists.

Methods

Study Framework

We followed the scoping review framework developed by Arksey and O'Malley [11], and conducted the reporting using the PRISMA Extension for Scoping Reviews (PRISMA-ScR) Checklist [12]. We carried out the following 5 stages of a scoping review: (1) identify the research question, (2) identify relevant studies, (3) select articles, (4) chart the data, and (5) collate and summarize the data [13]. To build the search strategy, we used the SPIDER tool (sample, phenomenon of interest, design, evaluation, research type) to identify qualitative and mixed method studies [14]. We also used the traditional PICO tool (patient, intervention, comparator, outcome) to develop a search strategy for quantitative studies, such as randomized controlled trials [15].

Information Sources

We searched the following databases for journal articles and conference proceedings between December 2018 and March 2019, and ran an update in January 2019: PubMed, Institute of Electrical and Electronics Engineers (IEEE), Association for Computing Machinery (ACM), International Pharmaceutical Abstracts (IPA), and EMBASE. Searches were conducted between December 2018 and January 2019. We also hand-searched reference lists from relevant articles. We exported all search results to EndNote reference manager software (version 8; Clarivate Analytics) and removed duplicates. The EndNote File was exported to Covidence (Veritas Health Innovation Ltd.).

Search

Three librarians worked together to build a comprehensive search strategy for each database, with support from database specialists. We began by familiarizing ourselves with the terminology for "reason for use" by conducting a preliminary search on PubMed and by searching reference lists of known publications on the topic. Developing a search strategy for each database was complex, balancing the need to be as comprehensive as possible while limiting the noise caused by the wide-reaching "indication" search term. Detailed search strategies are presented in Multimedia Appendix 1. A sample search strategy for PubMed is as follows:

((“reason for use”[All Fields] OR Indication*[All Fields] OR Off-Label Use[MeSH terms] OR (diagnosis[All Fields] OR diagnosis[MeSH terms] AND (pharmacists[MeSH Terms] OR

pharmacist*[All Fields])) AND (prescription[All Fields] OR drug prescriptions[MeSH Terms] OR prescriptions[MeSH Terms]) AND (documentation[MeSH Terms] OR document[All Fields] OR record[All Fields] OR communication [MeSH terms] OR communication[All Fields] OR Electronic health record[MeSH Terms] OR "electronic medical record" OR labels[All Fields] OR off-label[All Fields] OR Off-Label Use[MeSH Terms] OR electronic prescribing[MeSH Terms]) AND (collaboration OR intersectoral collaboration[MeSH Terms] OR interprofessional relations[MeSH Terms] OR patient care team[MeSH Terms] OR professional role[MeSH Terms] OR team[All Fields] OR interprofessional[All Fields] OR "interprofessional collaboration" [All Fields] OR patient[All Terms] OR patients[MeSH Terms]))).

Selection of Sources

We imputed titles and abstracts into Covidence and 2 authors (KM and CC) independently screened the titles, abstracts, and full-text articles according to the eligibility criteria. Studies were eligible for inclusion if they included pharmacists as part of the study and examined one of the following: (1) the inclusion of reason for use in a prescription; (2) the addition of reason for use to a prescription medication label; or (3) why prescribers do or do not include reason for use in prescriptions. We did not limit ourselves to a specific type of study, or field of study. We did not place any limits on the date or location of publications other than the research must be published in English. We excluded dissertations and commentaries.

Data Synthesis

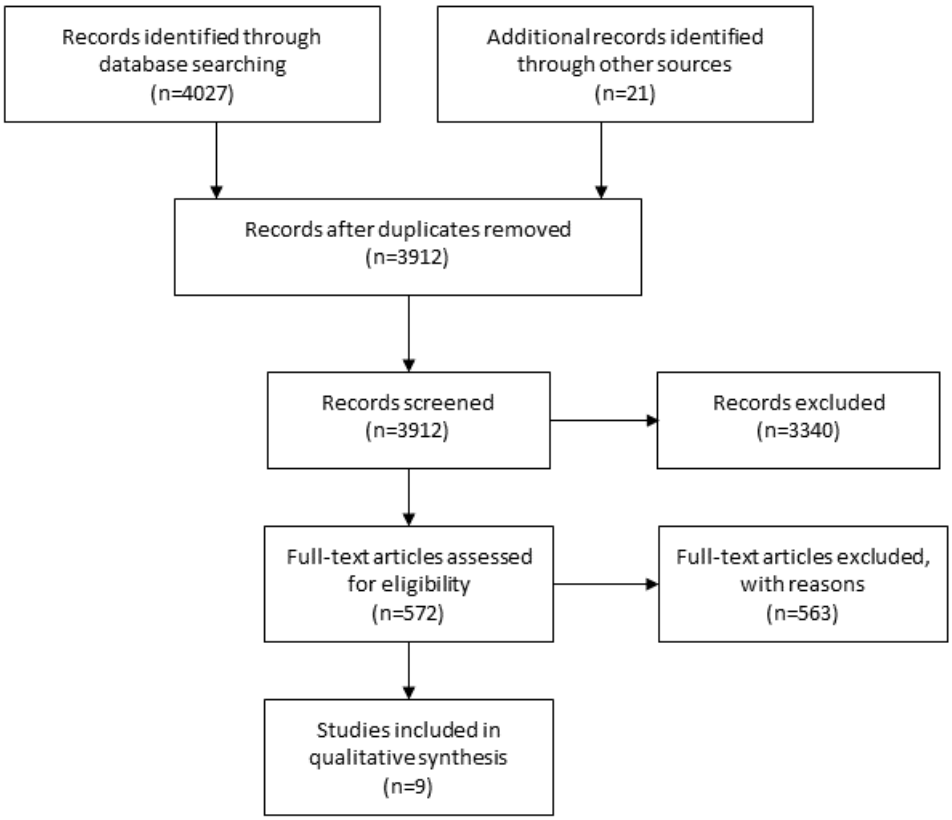
One researcher used a standardized form to extract data from included full-text articles, and the data were verified by a second researcher. We recorded the following data: lead author, year of publication, geographic location, participants, methods, analysis, research setting, outcomes, and location of the reason for use (eg, electronic health record, written prescription). While reviewing the included articles, we were guided by the research question for this study: "How are pharmacists affected when the reason for use is included on a prescription, and what are its implications for collaboration and patient safety." We began by categorizing the literature according to the methodology, key findings, and setting. As we reviewed the articles, we added categories as necessary to understand the full extent of themes and research currently being carried out. We identified gaps and key findings after reviewing the final list of included articles.

Results

Study Selection

We identified a total of 4027 titles with an additional 21 studies identified from other sources (Figure 1), of which 136 were duplicates. After screening, 3912 articles were screened, leaving a total of 9 that met the inclusion criteria [2,16-23] (Multimedia Appendix 2). Examples of reasons papers were excluded included the following: focus on labeling not prescriptions [24,25], did not include a pharmacist [26-30], focused on medication review without indication [31,32], monitoring drug treatment [33], and network data mining [2,16-23,34].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) diagram.



The literature was synthesized into 4 areas of focus: (1) terminology (2) importance of including reason for use on prescriptions; (3) impact of reason for use on decision making and workflow; and (4) barriers to reason for use information.

Descriptive Characteristics

The 9 included studies were published between 1998 and 2018. Six studies examined pharmacists and physicians jointly in the studies [16-18,21-23]. In total, 4 studies were conducted in the United States [2,18,21,23], 2 in Europe [20,22], 1 in the Middle East [16], and 2 in Australia [17,19]. Two studies focused on prescribing in hospital [16,22], 6 focused on primary care [2,17,19-21,23], and 1 involved a consultation with experts from different settings [18]. According to the inclusion criteria, all studies included pharmacists, with 7 also including physicians [16-19,21-23], 4 included patients [18,21-23], and 1 presented results from a pilot study with various stakeholders [18]. Five studies used a qualitative approach to capture perspectives on the application of reason for use [17,18,20,21,23] and 4 used a quantitative approach to characterize the impact of reason for use [2,16,19,22]. Three of the included studies were published in health research journals [17,19,22], with the remaining 6 published in pharmacy practice journals [2,16,18,20,21,23]. We did not identify any studies in the engineering or informatics literature.

Terminology

Including a reason for use on a prescription was described in a variety of ways. The most common terminology is “indication” [17,18,20] or related terms including “indication in prescription” [16], “medication indication [21], and “indication for treatment” [22]. Other terminologies were patient diagnosis [23], “reason for use” [2], purpose of the medication” [19], and “clinical patient data” [35].

Current Perspectives on Including Reason for Use on Prescriptions

All included studies identified that reason for use is needed to improve patient safety. Generally, the pharmacist and physician research participants had positive reactions toward adding the reason for use to prescriptions. Using semistructured interviews with pharmacists, physicians, and patients, Garada et al [17] identified that the addition of reason for use information can reduce perceived prescribing and dispensing errors, and that adding the information to the label supports patient engagement and the work of other health care professionals. Liddell and Goldman [19] specifically identified that including the reason for general use was the most important aspect of new prescription notations to improve communication.

Impact of Reason for Use on Decision Making and Workflow

Three studies mentioned pharmacists feeling limited by missing information [20,21,23], 3 identified the reason for use as being important to pharmacists for catching prescribing errors and improving safety [16,17,22], 4 recognized the potential for reason for use information to improve workflow [16,18,21,23], and 3 discussed the need for reason for use to provide accurate patient counseling [21,23,35]. Of the 3 studies that examined workflow, Al-Khani et al [16] identified the difficulty in getting physicians to comply with including reason for use, and the subsequent change in workflow.

Al-Khani et al’s [16] study used a hospital’s safety reporting system to show that 35% of the drug prescribing errors that pharmacists flagged were identified using reason for use. Liddell and Goldman [19] demonstrated a very positive response from physicians about being more collaborative with pharmacists when notations were included that specify the reason for use information, and both pharmacists and physicians were positive about tools that would facilitate communication.

Improved collaboration and communication between pharmacists and physicians were identified in 2 articles [20,21]. Tarn et al [21] identified the potential benefit that improved collaboration can have on efficiency. Kron et al [18] discussed how pharmacists often try to infer information about why a medication was prescribed from the patients, which is supported by Warholak et al’s [23] findings that after a diagnosis was included on an electronic prescription, pharmacists have less confusion and uncertainty [18,23], further identifying that patients are used as an intermediary to get access to information.

Barriers to the Reason for Use Information

Only 1 paper examined privacy concerns, concluding that while pharmacists and physicians were concerned about privacy, patients were not generally concerned with the privacy implications of documenting reason for use on a prescription [17]. Of the 5 included studies that mentioned technology [16-18,22,23], 4 suggested there was a need to improve the prescribing software available [17,18,22,23]. Four studies examined electronic prescribing [16-18,23].

Raebel et al [35] discussed the effectiveness of a computerized pharmacy alert system and active collaboration between health care professionals. The study’s goal was to improve prescribing safety and identified that a barrier to this was that clinical patient data were not easily available to many pharmacists [35]. Kron et al [18] specifically examined the difficulties in encouraging prescribers to include the reason for the prescription, and identified that electronic prescribing was laying the foundation for future adoption.

Discussion

Principal Findings

We set out to identify and describe the current literature around how the reason for use information can be shared with a pharmacist through a prescription. We identified several studies where systems supported a mandatory reason for use field or modified the computer interface to make it easier for prescribers to add this information. The research to date has not moved much beyond the typical barriers to adoption such as a lack of time or incentives; however, when asked, prescribers do clearly identify the benefits of adding this information. Therefore, one of the key findings of this review is we did not identify any implementation science or theory-based studies aiming to improve adoption by prescribers.

One observation from this study was that the problem of sharing reason for use information is greater than making the reason for use field mandatory for prescribers. The study by Al-Khani et al [16] in a hospital in Saudi Arabia retrospectively identified that pharmacists' access to the reason for use and medication history was a major factor in identifying up to 60% of errors, even though some physicians found ways to override the mandatory field on prescriptions by writing characters or letters. Through their experiences in a Chinese hospital, Li and Zhou [36] also highlighted that a hospital-wide policy could promote the addition of reason for use to electronic prescriptions—allowing the pharmacy department to keep the proportion of "inappropriate physician orders" below 1%—but that it raised a new challenge when prescribers provide poor quality or incomplete information [36]. Another study in a Dutch children's hospital also found that prescribers rejected up to half of pharmacist recommendations due to a lack of timeliness or relevance—highlighting that improved information around indication would lead to more timely and meaningful collaboration between prescribers and pharmacists [37]. Moving forward, researchers could benefit from using an implementation science framework to look at the reasons interventions or policies worked (or not), especially in critical areas related to reach, effectiveness, adoption, implementation, and maintenance over the long term [38].

One major limitation of this review was that the terminology used to describe "reason for use" was not consistent, which is a barrier to building an evidence-based body of knowledge to encourage designing, implementing, and ultimately having an uptake of including a reason for use on prescriptions. If the language used is not consistent, it is difficult to make sure all stakeholders are working on the same problem, toward the same solution. The reason for use literature bridges health, engineering, informatics, and other areas, all with different terminologies, frameworks, and methods. The papers included in this study were all from health care journals, primarily pharmacy journals. This may mean that engineering and informatics disciplines are not aware of these papers. While the methodology for health-related scoping reviews is well documented [11], the search methodology and available tools have not yet caught up in other disciplines. For example, while PubMed uses the MeSH search terms and EMBASE uses Emtree, these are not standard between databases, and the nonmedical databases do not have standardized search terms.

Conclusions

In the limited literature to date, there is a consensus that the addition of reason for use information to prescriptions benefits patient safety and enables pharmacists to be more effective. However, it is also clear that very little has been done to motivate prescribers to include this information, despite clear benefits such as reducing the number of phone calls received by pharmacists. Future research should be multidisciplinary, and use an implementation science or theory-based approach to improve prescriber buy-in and, consequently, adoption.

Acknowledgments

We acknowledge the expertise and guidance of Jackie Stapleton, a health librarian at the University of Waterloo, who helped with the development and refinement of the search strategies used in this review. We also thank the assistance provided by Cassandra Sum and Parvinder Sahota.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Final searches.
[PDF File (Adobe PDF File), 92 KB - humanfactors_v8i4e22325_app1.pdf]

Multimedia Appendix 2
Included studies.
[PDF File (Adobe PDF File), 84 KB - humanfactors_v8i4e22325_app2.pdf]

References

1. Garada M, McLachlan AJ, Schiff GD, Lehnbom EC. What do Australian consumers, pharmacists and prescribers think about documenting indications on prescriptions and dispensed medicines labels?: A qualitative study. BMC Health Serv Res 2017 Nov 15;17(1):734 [FREE Full text] [doi: 10.1186/s12913-017-2704-3] [Medline: 29141618]
2. Warholak-Juarez T, Rupp MT, Salazar TA, Foster S. Effect of patient information on the quality of pharmacists' drug use review decisions. J Am Pharm Assoc (Wash) 2000;40(4):500-508. [Medline: 10932459]
3. Mercer K, Burns C, Guirguis L, Chin J, Dogba MJ, Dolovich L, et al. Physician and Pharmacist Medication Decision-Making in the Time of Electronic Health Records: Mixed-Methods Study. JMIR Hum Factors 2018 Sep 25;5(3):e24 [FREE Full text] [doi: 10.2196/humanfactors.9891] [Medline: 30274959]

4. Formica MK, McAlindon TE, Lash TL, Demissie S, Rosenberg L. Validity of self-reported rheumatoid arthritis in a large cohort: results from the Black Women's Health Study. Arthritis Care Res (Hoboken) 2010 Feb;62(2):235-241 [FREE Full text] [doi: 10.1002/acr.20073] [Medline: 20191523]
5. Gure TR, McCammon RJ, Cigolle CT, Koelling TM, Blaum CS, Langa KM. Predictors of Self-Report of Heart Failure in a Population-Based Survey of Older Adults. Circ Cardiovasc Qual Outcomes 2012 May;5(3):396-402. [doi: 10.1161/circoutcomes.111.963116]
6. Schneider ALC, Pankow JS, Heiss G, Selvin E. Validity and reliability of self-reported diabetes in the Atherosclerosis Risk in Communities Study. Am J Epidemiol 2012 Oct 15;176(8):738-743 [FREE Full text] [doi: 10.1093/aje/kws156] [Medline: 23013620]
7. Adler-Milstein J, Holmgren A, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital "advanced use" divide. J Am Med Inform Assoc 2017 Nov 01;24(6):1142-1148 [FREE Full text] [doi: 10.1093/jamia/ocx080] [Medline: 29016973]
8. Mooranian A, Emmerton L, Hattingh L. The introduction of the national e-health record into Australian community pharmacy practice: pharmacists' perceptions. Int J Pharm Pract 2013 Dec;21(6):405-412. [doi: 10.1111/ijpp.12034] [Medline: 23560554]
9. Jackson S, Peterson G. My Health Record: a community pharmacy perspective. Aust Prescr 2019 Apr;42(2):46-47 [FREE Full text] [doi: 10.18773/austprescr.2019.009] [Medline: 31048935]
10. Schiff GD, Seoane-Vazquez E, Wright A. Incorporating Indications into Medication Ordering — Time to Enter the Age of Reason. N Engl J Med 2016 Jul 28;375(4):306-309. [doi: 10.1056/nejmp1603964]
11. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
12. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
13. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
14. Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. Qual Health Res 2012 Oct;22(10):1435-1443. [doi: 10.1177/1049732312452938] [Medline: 22829486]
15. Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. Qual Health Res 2012 Oct;22(10):1435-1443. [doi: 10.1177/1049732312452938] [Medline: 22829486]
16. Al-Khani S, Moharram A, Aljadhey H. Factors contributing to the identification and prevention of incorrect drug prescribing errors in outpatient setting. Saudi Pharm J 2014 Nov;22(5):429-432 [FREE Full text] [doi: 10.1016/j.jsps.2013.11.003] [Medline: 25473331]
17. Garada M, McLachlan AJ, Schiff GD, Lehnbom EC. What do Australian consumers, pharmacists and prescribers think about documenting indications on prescriptions and dispensed medicines labels?: A qualitative study. BMC Health Serv Res 2017 Nov 15;17(1):734 [FREE Full text] [doi: 10.1186/s12913-017-2704-3] [Medline: 29141618]
18. Kron K, Myers S, Volk L, Nathan A, Neri P, Salazar A, et al. Incorporating medication indications into the prescribing process. Am J Health Syst Pharm 2018 Jun 01;75(11):774-783. [doi: 10.2146/ajhp170346] [Medline: 29674327]
19. Liddell M, Goldman S. Attitudes to and use of a modified prescription form by general practitioners and pharmacists. Med J Aust 1998 Apr 06;168(7):322-325. [doi: 10.5694/j.1326-5377.1998.tb138959.x] [Medline: 9577441]
20. Stakenborg JPG, de Bont EGPM, Peetoom KKB, Nelissen-Vrancken MHJMG, Cals JWL. Medication management of febrile children: a qualitative study on pharmacy employees' experiences. Int J Clin Pharm 2016 Oct;38(5):1200-1209 [FREE Full text] [doi: 10.1007/s11096-016-0353-y] [Medline: 27450505]
21. Tarn D, Paterniti D, Wenger N, Williams B, Chewning B. Older patient, physician and pharmacist perspectives about community pharmacists' roles. Int J Pharm Pract 2012 Oct;20(5):285-293 [FREE Full text] [doi: 10.1111/j.2042-7174.2012.00202.x] [Medline: 22953767]
22. Vercheval C, Gillet M, Maes N, Albert A, Frippiat F, Damas P, et al. Quality of documentation on antibiotic therapy in medical records: evaluation of combined interventions in a teaching hospital by repeated point prevalence survey. Eur J Clin Microbiol Infect Dis 2016 Sep;35(9):1495-1500. [doi: 10.1007/s10096-016-2690-0] [Medline: 27255220]
23. Warholak TL, Rupp MT, Leal S, Kurniawan G, Patel N. Assessing the effect of providing a pharmacist with patient diagnosis on electronic prescription orders: a pilot study. Res Social Adm Pharm 2014;10(1):246-251. [doi: 10.1016/j.sapharm.2013.04.010] [Medline: 23688539]
24. Levy G, Zamacona M, Jusko W. Developing compliance instructions for drug labeling. Clin Pharmacol Ther 2000 Dec;68(6):586-591. [doi: 10.1067/mcp.2000.110976] [Medline: 11180017]
25. Zargarzadeh AH, Law AV. Design and test of preference for a new prescription medication label. Int J Clin Pharm 2011 Apr;33(2):252-259. [doi: 10.1007/s11096-011-9488-z] [Medline: 21394572]
26. Murphy O, King G, Callanan IW, Butler MW. Why is your patient sedated? Int J Health Care Qual Assur 2014;27(4):284-292. [doi: 10.1108/IJHCQA-06-2013-0074] [Medline: 25076603]

27. Eguale T, Winslade N, Hanley JA, Buckeridge DL, Tamblyn R. Enhancing pharmacosurveillance with systematic collection of treatment indication in electronic prescribing: a validation study in Canada. Drug Saf 2010 Jul 01;33(7):559-567. [doi: 10.2165/11534580-000000000-00000] [Medline: 20553057]
28. Curtis LM, Mullen RJ, Russell A, Fata A, Bailey SC, Makoul G, et al. An efficacy trial of an electronic health record-based strategy to inform patients on safe medication use: The role of written and spoken communication. Patient Educ Couns 2016 Sep;99(9):1489-1495 [FREE Full text] [doi: 10.1016/j.pec.2016.07.004] [Medline: 27444235]
29. Kuo GM, Mullen PD, McQueen A, Swank PR, Rogers JC. Cross-sectional comparison of electronic and paper medical records on medication counseling in primary care clinics: a Southern Primary-Care Urban Research Network (SPUR-Net) study. J Am Board Fam Med 2007;20(2):164-173 [FREE Full text] [doi: 10.3122/jabfm.2007.02.060113] [Medline: 17341753]
30. Holsappel IGA, Koster ES, Winters NA, Bouvy ML. Prescribing with indication: uptake of regulations in current practice and patients opinions in the Netherlands. Int J Clin Pharm 2014 Apr;36(2):282-286. [doi: 10.1007/s11096-013-9903-8] [Medline: 24337662]
31. Nusair MB, Guirguis LM. How pharmacists check the appropriateness of drug therapy? Observations in community pharmacy. Res Social Adm Pharm 2017;13(2):349-357. [doi: 10.1016/j.sapharm.2016.03.004] [Medline: 27102265]
32. Slee A, Farrar K, Hughes D, Constable S. Optimising Medical Treatment-How Pharmacist-Acquired Medication Histories Have a Positive Impact on Patient Care. Vol 277.. Accessed March 14. www.pjonline.com 2006:737-739.
33. Goethe J, Bronzino J. An expert system for monitoring psychiatric treatment. IEEE Eng. Med. Biol. Mag 2005 Nov;14(6):776-780. [doi: 10.1109/51.473274]
34. Zhao M. Off-Label Drug Use Detection Based on Heterogeneous Network Mining. : IEEE; 2017 Aug Presented at: InIEEE International Conference on Healthcare Informatics (ICHI). IEEE; 2017; Park City, UT, USA p. 2017 URL: https://ieeexplore.ieee.org/xpl/conhome/8030514/proceeding [doi: 10.1109/ichi.2017.33]
35. Raebel M, Charles J, Dugan J, Carroll NM, Korner EJ, Brand DW, et al. Randomized trial to improve prescribing safety in ambulatory elderly patients. J Am Geriatr Soc 2007 Jul;55(7):977-985. [doi: 10.1111/j.1532-5415.2007.01202.x] [Medline: 17608868]
36. Li W, Zhou Q. Patient diagnosis on electronic prescription orders: a key element for appropriateness review by pharmacists. Res Social Adm Pharm 2013;9(4):493-494. [doi: 10.1016/j.sapharm.2013.06.005] [Medline: 23890197]
37. Maat B, Au Y, Bollen C, van Vught AJ, Egberts TCG, Rademaker CMA. Clinical pharmacy interventions in paediatric electronic prescriptions. Arch Dis Child 2013 Mar;98(3):222-227. [doi: 10.1136/archdischild-2012-302817] [Medline: 23184352]
38. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM Planning and Evaluation Framework: Adapting to New Science and Practice With a 20-Year Review. Front Public Health 2019;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]

Edited by R Kukafka, G Eysenbach; submitted 08.07.20; peer-reviewed by K Fuji, N Finn; comments to author 22.08.20; revised version received 17.11.20; accepted 02.08.21; published 25.11.21.

Please cite as:
Mercer K, Carter C, Burns C, Tennant R, Guirguis L, Grindrod K
Including the Reason for Use on Prescriptions Sent to Pharmacists: Scoping Review
JMIR Hum Factors 2021;8(4):e22325
URL: https://humanfactors.jmir.org/2021/4/e22325
doi:10.2196/22325
PMID:34842545

Original Paper

A Novel Digital Pill System for Medication Adherence Measurement and Reporting: Usability Validation Study

Susan L Baumgartner[1*], PharmD, MBA; D Eric Buffkin Jr[1*], BSEE, MBA; Elise Rukavina[2*], MSc; Jason Jones[2*], BSc, MSc; Elizabeth Weiler[2*]; Tony C Carnes[1*], PhD

[1]etectRx, Inc., Gainesville, FL, United States

[2]Tensentric, Inc., Boulder, CO, United States

[*]all authors contributed equally

Corresponding Author:
Susan L Baumgartner, PharmD, MBA
etectRx, Inc.
747 SW 2nd Avenue
Suite 365T, IMB 24
Gainesville, FL, 32601
United States
Phone: 1 678 602 5701
Email: susan.baumgartner@etectrx.com

Abstract

Background: Medication nonadherence is a costly problem that is common in clinical use and clinical trials alike, with significant adverse consequences. Digital pill systems have proved to be effective and safe solutions to the challenges of nonadherence, with documented success in improving adherence and health outcomes.

Objective: The aim of this human factors validation study is to evaluate a novel digital pill system, the ID-Cap System from etectRx, for usability among patient users in a simulated real-world use environment.

Methods: A total of 17 patients with diverse backgrounds who regularly take oral prescription medications were recruited. After training and a period of training decay, the participants were asked to complete 12 patient-use scenarios during which errors or difficulties were logged. The participants were also interviewed about their experiences with the ID-Cap System.

Results: The participants ranged in age from 27 to 74 years (mean 51 years, SD 13.8 years), and they were heterogeneous in other demographic factors as well, such as education level, handedness, and sex. In this human factors validation study, the patient users completed 97.5% (196/201) of the total use scenarios successfully; 75.1% (151/201) were completed without any failures or errors. The participants found the ID-Cap System easy to use, and they were able to accurately and proficiently record ingestion events using the device.

Conclusions: The participants demonstrated the ability to safely and effectively use the ID-Cap System for its intended use. The ID-Cap System has great potential as a useful tool for encouraging medication adherence and can be easily implemented by patient users.

(JMIR Hum Factors 2021;8(4):e30786) doi:10.2196/30786

KEYWORDS

digital pills; digital medication; ingestible event marker; ingestible sensor; human factors; usability; validation study; medication adherence; medication nonadherence; remote patient monitoring; mobile phone

Introduction

Background

Medication nonadherence is a problem that continues to plague the health care system. Data have shown that patients do not report their own adherence accurately [1] and that health care providers are generally poor judges of their patients' adherence [2]. In clinical practice, it is known that up to 50% of patients do not take their medications as prescribed [3], even in serious disease states or conditions where the consequences can be severe, such as diabetes, heart failure, hyperlipidemia, hypertension, and organ transplantation [4,5].

In addition, medications that have been shown to improve the quality of life, prevent tumor progression, and prolong survival are often not taken as prescribed by patients with cancer [6]. A systematic review of publications on oral anticancer medications from 2003 to 2015 showed that medication adherence rates varied widely from 46% to 100% [7]. In interviews with patients with breast cancer, de Mendoza et al [8] found that 78.9% of the patients failed to report medication discontinuation immediately and 57.9% overreported medication adherence.

Nonadherence is multifactorial. The common reasons for nonadherence are confusion (about complex drug regimens), a lack of commitment to the treatment plan, fear of adverse events, cost of drugs, forgetfulness, lack of symptoms, illness factors such as depression or psychosis, and miscommunication or lack of trust between the patient and the health care team [9-11].

Clinical trials too are often impaired by suboptimal adherence and flawed in the way they track medication adherence [12-14]. For example, in a systematic review that captured adherence data from 95 clinical trials involving 16,907 participants, there was an immediate 4% drop-off of the enrolled participants because of noninitiation of therapy. By day 100, 20% of the participants had stopped taking the medication. A further 12% displayed imperfect adherence on a daily basis [14]. Adherence errors can result in suboptimal dosing and inaccurate assessments of efficacy, safety, and tolerability, thus delaying the drug development process and potentially adding millions of dollars in additional costs [9,15].

The need for objective and reliable ways to confirm medication use has driven the development of various tracking methods, including patient self-reports, adherence-reporting mobile apps, pill counts, pharmacy prescription refill rates, electronic pill dispensers, and other solutions—but none have been optimized, and many are not reliable [16]. Digital pill systems, in contrast, have demonstrated a high rate of accuracy, with a study showing a 99.4% adherence rate across 2824 digital pill ingestions that were tracked [17].

The ID-Cap System (etectRx, Inc) is a digital pill system and ingestible event marker (Code of Federal Regulations 21 §880.6305) that enables adherence measurement through an embedded ingestible sensor. The biocompatible sensor, upon coming into contact with gastrointestinal fluid, communicates a digital signal through radio frequency after ingestion and dissolution of the pharmaceutical-grade capsule shell that encapsulates it. A reader worn by the patient detects the radio frequency signal and forwards ingestion data to the patient app and clinician dashboard. Information about the ingestion event is then wirelessly transferred to a Health Insurance Portability and Accountability Act–compliant cloud-based server [18] for secure sharing with authorized users. The sensor is naturally and safely eliminated from the body.

The Food and Drug Administration granted 510(k) clearance to the ID-Cap System in December 2019. The regulatory review of this medical device and its related software encompassed the results of human factors validation testing among patient users, clinician users, and system administrators, including the results of the study reported herein. The components of the ID-Cap System are shown in Figures 1-5. The patient user testing was conducted with the following system components: the ID-Capsule, the ID-Cap Reader, and the ID-Cap Patient App.

Figure 1. ID-Capsule: a digital pill consisting of a pharmaceutical-grade capsule shell with an embedded ingestible sensor. The ID-Capsule has been designed to encapsulate medications that are tracked using the system. The sensor communicates a digital signal shortly after ingestion and capsule dissolution. The sensor is naturally and safely eliminated through the patient's gastrointestinal tract.

Figure 2. ID-Cap Reader: a wearable device that detects messages transmitted from the ingested sensor and forwards them to the ID-Cap Patient App and Clinician Dashboard.

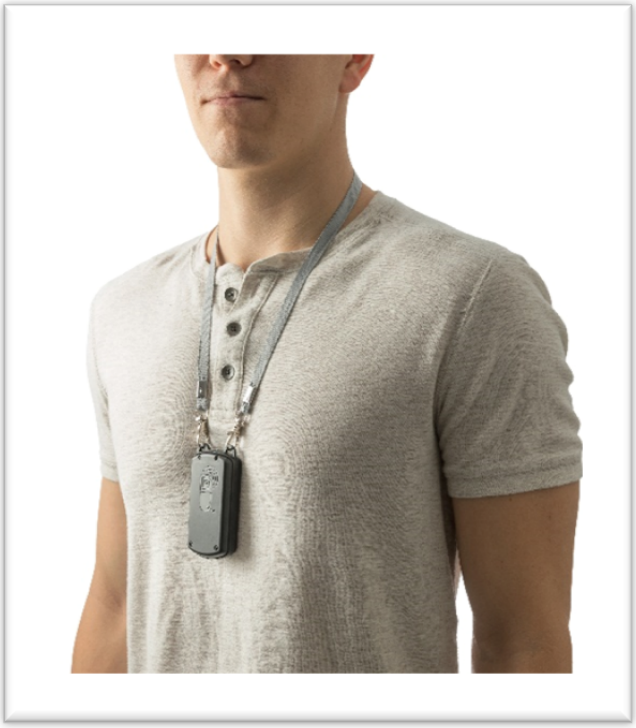


Figure 3. The ID-Cap Patient App allows patients to view ingestion events in real time as well as their medication use history. The app can also send patient reminders and alerts.

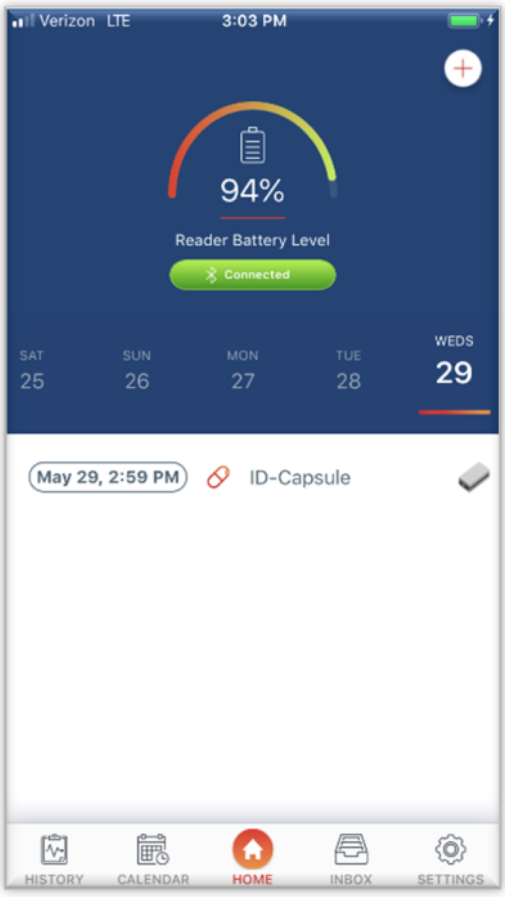


Figure 4. The Clinician Dashboard enables logging, tracking, and trending of patients' ingestion events by clinicians. It provides both real-time notifications and a history of ingestion events.

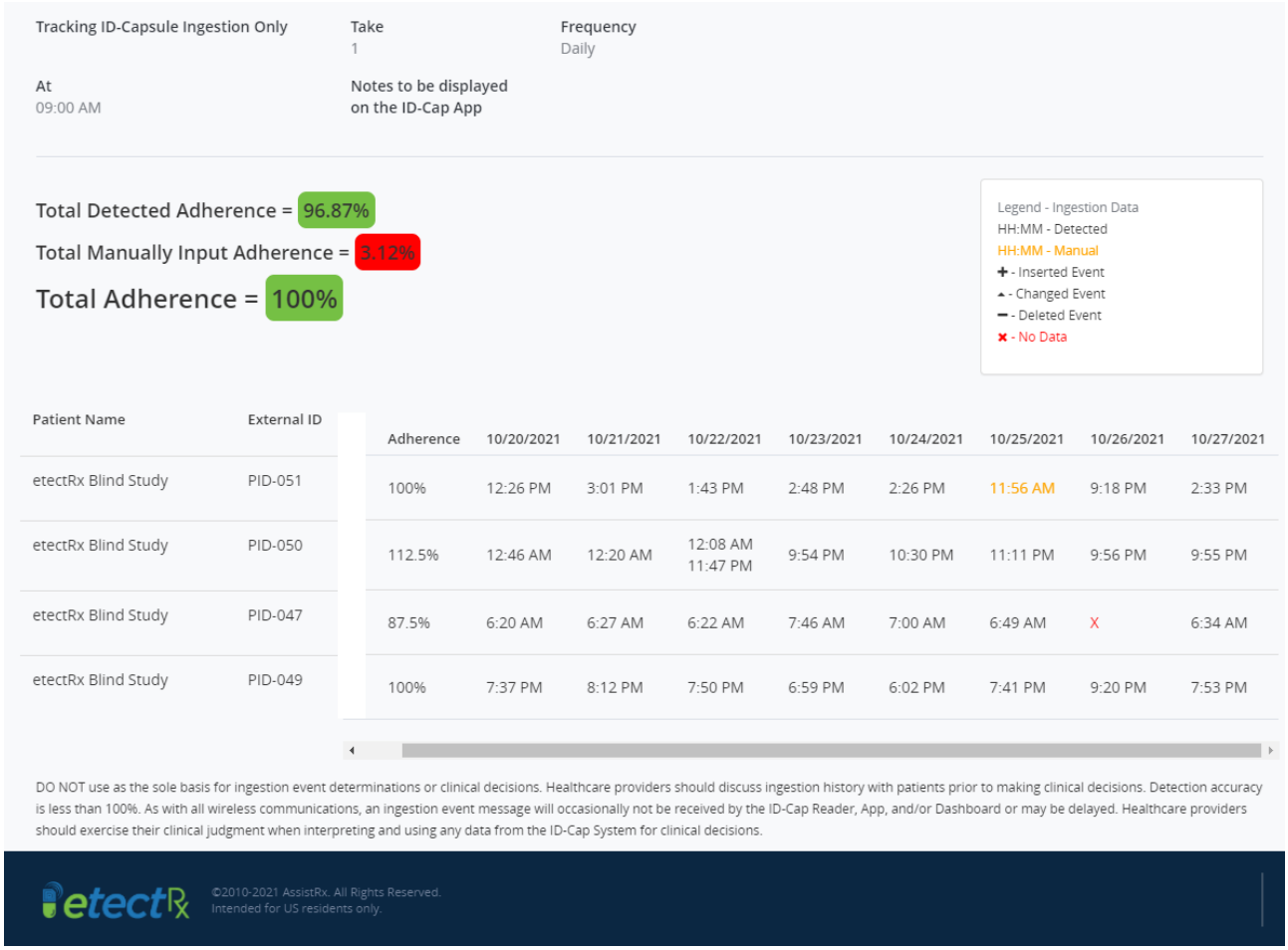
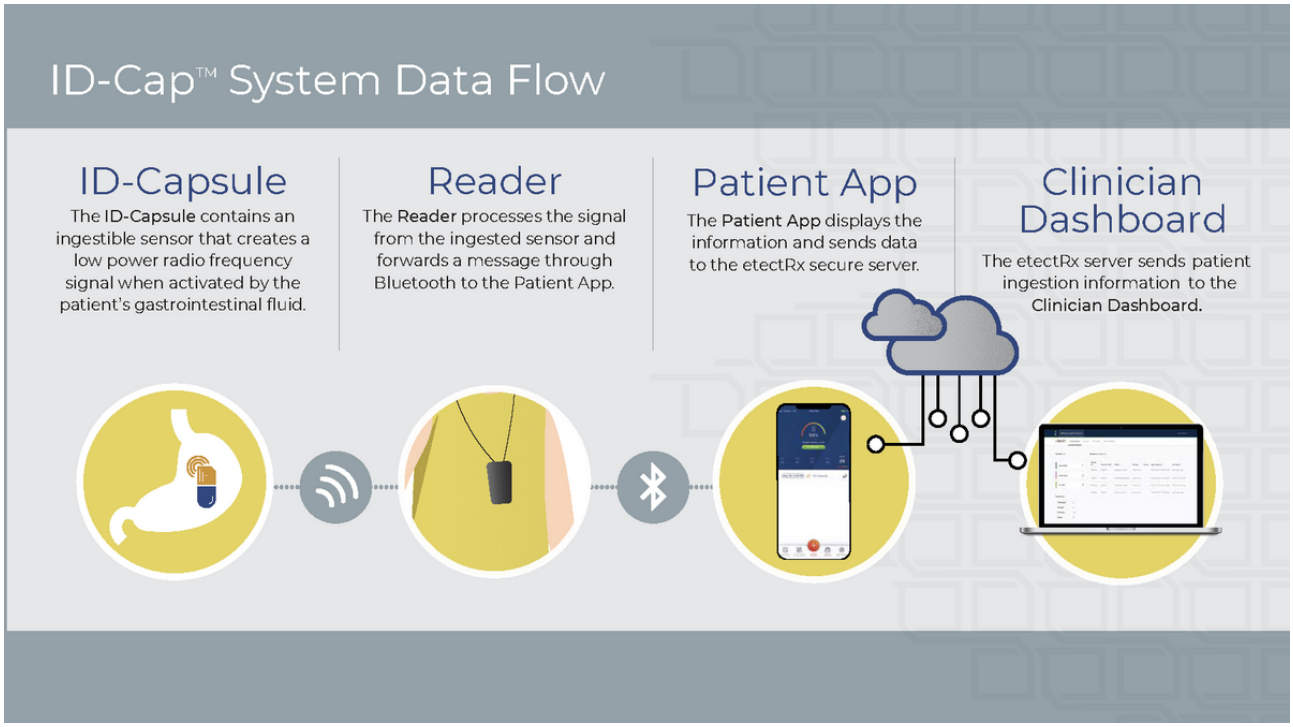


Figure 5. Overview of how the ID-Cap System works.



Objective

In this paper, we describe the human factors validation study involving patient users, the intent of which is to evaluate the ID-Cap System for usability, ensuring that patient users will be able to operate the system as intended in a simulated real-world use environment. Our key questions were as follows:

1. Are patient users able to perform critical tasks effectively and safely, using an interface representative of the final device design, in conditions representing the actual conditions of use?
2. Can they do so without errors and without difficulties that could cause harm?

This validation study excluded clinical safety and effectiveness elements, which have been assessed and documented in separate evaluations and pivotal clinical trials supporting the use of the device. The researchers hypothesized a priori that the patient users would successfully demonstrate their ability to safely and effectively use the ID-Cap System for its intended use.

Methods

Overview

To conduct the human factors validation test of the ID-Cap System, we used assessment testing. This type of test provides users with realistic tasks to perform, using a working prototype of the device but without requiring any clinical use [19].

The training and testing took place at the office facilities of Tensentric, Inc, in Boulder, Colorado. The test was conducted in either conference rooms or dedicated research rooms set up to represent a typical home-use environment. An ID-Cap System was provided, including supplemental test equipment (eg, laptops for training videos) and product labeling.

Participants

The research team had a recruitment target of up to 18 patient users to ensure that the goal of a minimum of 15 participants from the intended user population would be met; 15 test participants per user group represents the most stringent sample size guidance from regulatory bodies for human factors validation testing [20]. The participants were recruited by an independent third-party recruiting firm that had no knowledge of etectRx's involvement at the time of recruitment. None of the participants were employed by, or affiliated with, etectRx, nor had they participated in a preceding formative usability or validation test of the ID-Cap System. Each participant signed a nondisclosure agreement and an informed consent form documenting their agreement for participation in the test session and video recording. Each participant received an honorarium for participation in the study, which was distributed after the completion of the test session.

Our goal was to obtain a sample of test participants who represent the intended patient users of the ID-Cap System. The recruited participants took medication by mouth on a regular basis, were able to understand and follow directions, and communicated clearly. If a caregiver assisted the participant in taking medication or in day-to-day activities, the caregiver also participated in the test. Participants with conditions that affected their ability to make health decisions or follow their physician's instructions and those who did not use a smartphone were excluded.

This was an all-comers study that was, by design, both inclusive and diverse. Diversity within the sample size was promoted within the recruitment screener by quantified maximums or minimums on specific populations as well as instructions to recruit a variety of participants, including a mix of handedness, sexes, races, educational levels, and disease states. The demographics that were recorded after scheduling included visual acuity, age, handedness, sex, medical conditions, how the participant currently ensures that they are taking their medication, color blindness, visual impairments, and any visual corrections.

Training

The participants received a 40-minute training session for the ID-Cap System to orient them to the basic features, functions, and nomenclature of the ingestible sensor, wearable reader, and patient app. They were shown a series of 5 short training videos containing key information and demonstrating the correct use of the device. They were then asked to demonstrate the steps for completing a successful ingestion using the reader and patient app before using the system independently during the test session. The training was equivalent to what is expected to be delivered to actual users, and the content, format, and method of delivery of training were comparable to the training that actual users would receive. The training materials and device instructions for use were designed to support a self-guided supplemental training program for the patient user.

The training videos covered the following topics:

- Overview of the ID-Cap System
- Setup of charger and reader for the first time
- Routine use of the system *without* use of the patient app
- Routine use of the system *with* use of the patient app
- App navigation and functionality

Training was conducted at least one hour before the patient's test session to simulate a typical level of training decay. After completing the training, the participants were sent away from the test environment for at least one hour with no materials. Only after the waiting period of an hour were the participants able to begin executing the test scenarios. This 1-hour gap, which was added to be more indicative of real-world gaps that exist between training and first-time use, represents the recommended time frame when evaluating potential use-related risks related to training decay [19].

Testing

After the training session and training decay period, the participants initiated a guided 60-minute test session. Time variation was expected in completion of the test session based on factors such as operator skill, training retention, and experience with similar devices and smartphone apps. Actual task times were documented through the video recording of each session and referenced as needed.

During the recorded test session, the participants were observed as they completed the use tasks and monitored by at least one

facilitator and either a second facilitator or an observer seated in the same room as the participants. Printed instructions for use were provided to the participants and placed where they would be freely available for reference. However, the facilitator did not direct the participants to use them; this enabled insight into how patients might (or might not) refer to them during actual use.

During the test session, the participants were asked to complete 12 patient-use scenarios using the ID-Cap System, which included device use tasks and knowledge assessment tasks. They verbally read the use-scenario description that was provided on a printed note card and started the task. The participants were instructed by the facilitator to use the system as independently and naturally as possible to reflect actual use behavior. They were not informed how to complete the task or when to expect error conditions nor were they given any other information that would bias their realistic interaction with the system.

The research team observed the participants as they attempted to carry out the task scenarios and recorded any difficulties or participant comments, which were revisited with the participant during a postscenario interview. A data logger observed and logged participant behavior, user comments, and system activity. The participant was asked to simulate the ingestion of the ID-Capsule and the medication that it was intended to track during the testing. No actual ingestion events occurred during the test session.

After each use scenario, the research facilitators conducted a postscenario interview with each participant to analyze any use-related problems. The participants were prompted to provide subjective and candid assessments of any use issues experienced during the test, their probable causes, and impact.

After all task scenarios were attempted, the participants completed a postsession debriefing interview with the facilitator, where neutrally worded, open-ended questions were posed to them regarding their experience. The participants were asked to provide feedback regarding the use, safety, and usability of the ID-Cap System, as well as the clarity and effectiveness of user resources containing instructions for its use.

Critical Tasks and Use Scenarios

This test protocol was designed to validate the critical use tasks associated with the ID-Cap System for patient users. The tasks were selected and prioritized using a risk-based analysis to cover critical tasks, safety-related tasks, frequent tasks, tasks that must be performed correctly for the device to work as intended, and key device labeling (Table 1).

The tasks that tested safety mitigations were given the highest priority; for example, tasks that have the potential to alter decisions about ID-Capsule ingestion or ingestions of medications that are taken coincident with, or co-ingested with, the ID-Capsule were determined to be the most important to evaluate. Next on the priority scale were tasks that enable proper operation of the system, followed by tasks that occur infrequently or are provided as a convenience to the user.

Table 1. Critical patient tasks were performed in various scenarios with the test participants, which allowed usability validation and risk assessment of the critical use tasks associated with the ID-Cap System.

Task ID	Task description
PT[a]01	Understand key device labeling for patient users
PT02	Power on reader before use
PT03	Set up charging pad and charge reader
PT08	Download and set up app on smartphone
PT09	Pair reader with smartphone
PT04	View, understand and respond to reader indicator light
PT05	Wear reader appropriately to record ingestion event
PT06	Ingest ID-Capsule (alone or co-ingested with medication)
PT07	Wear reader for sufficient time to record ingestion event
PT10	Understand and respond to ingestion confirmations from reader and app
PT11	Understand and respond to reminder notifications from app appropriately; for example, by manually recording an ingestion event that the reader did not record
PT12	View and understand ingestion history in app: • App properly records a detected ingestion event • Interpret reader-detected ingestions and manually-recorded ingestions

[a]PT: patient task.

Data Analysis and Reporting

The cross-functional research team members performed a risk-based review of the human factors validation test results, including the participants' subjective assessments of any use errors, close calls, or operational difficulties that occurred during the test. Final pass or fail determination was made based on the risk-based review of the test results and if further design modifications were required to mitigate use-related errors.

The final report summarized the test results, which included evaluation of the actual versus expected task outcomes, subjective assessments, any specific use-related problems, and recommendations for resolution.

Results

Demographics and Baseline Characteristics

A total of 17 participants met the recruitment criteria and were enrolled in the validation test, fulfilling the minimum target of 15 participants from the intended user population of patient users. The participants ranged in age from 27 to 74 years (mean 51 years, SD 13.8 years), and they were heterogeneous in other demographic factors as well, such as education level, handedness, and sex (Table 2). Of the 17 participants, 7 (41%) were women, and nearly one-quarter of the participants reported high school as the highest level of education attained. All participants reported taking prescription medications by mouth on a regular basis and using a smartphone. This is consistent with the expected user population for the ID-Cap System.

Table 2. Demographic characteristics of the patient user group (N=17).

Participant ID	Age (years)	Education	Handedness	Sex	Type of smartphone used
P01	64	High school graduate	Right	Male	iPhone 6
P02	50	Professional degree	Right	Male	iPhone 10 R
P03	44	Master's degree	Right	Female	iPhone 7
P04[a]	41	High school graduate	Right	Male	Samsung Galaxy S10
P05	55	Professional degree	Right	Male	iPhone 6
P06	27	Bachelor's degree	Right	Female	iPhone 6
P07	61	Bachelor's degree	Right	Female	iPhone 8
P08	69	Professional degree	Left	Male	iPhone 7
P09	39	Bachelor's degree	Right	Female	Samsung Galaxy S9
P10	40	Bachelor's degree	Right	Female	iPhone 7
P11	50	High school graduate	Right	Male	Android
P12	74	Bachelor's degree	Right	Male	LG
P13	33	Professional degree	Right	Female	iPhone XR
P14	74	Bachelor's degree	Right	Male	iPhone 7
P15	48	Associate degree	Left	Male	Samsung Galaxy S9
P16	57	Doctorate degree	Right	Female	iPhone 8
P17	49	Some college—no degree	Left	Male	Samsung S7 Edge

[a]P04 was a wheelchair-bound quadriplegic person who was accompanied by a caregiver for the test session.

Test Session Results

In this validation study, the participants successfully completed 97.5% (196/201) of the total patient use scenarios with the ID-Cap System. Of the 12 use scenarios, 9 (75%) were successfully completed without any failures or use errors among the 17 test participants.

The task scenarios included first-time use tasks and repeat-use tasks, as well as tasks that were only completed on a single occasion, as appropriate, to represent actual use of the system. Of note, we found that when use errors did occur, they occurred only in the first instance and were not repeated. The use errors, close calls, and patterns of use difficulties identified in the testing are reported in Table 3.

The results of the human factors validation test were reviewed by a cross-functional team that conducted a risk-based review of each use-related finding. No new use-related risks were identified in the validation test. It was determined that no modifications of the device or software were required to improve safety or usability for the intended patient users, uses, and use environments of the ID-Cap System when operated as indicated and in a manner consistent with its labeling.

Table 3. Use errors, close calls, and patterns of use difficulty encountered within patient scenarios and patient knowledge tasks. Successful completion rate is the percentage of participants who successfully completed the task among those who attempted it (N=17).

Scenario	Title	Successful completion rate, % (number of participants who successfully completed the task/number of participants who attempted the task)	Summary of use errors, close calls, or use difficulties
PS[a]01	Set up & confirm reader is ready for use	100 (17/17)	• A few participants had difficulty turning on the reader, but all were ultimately able to do so after 1-2 minutes and without assistance from the test facilitator
PS02	Record ingestion event using the ID-Cap reader: ID-Capsule alone	88 (15/17)	• Several participants took the ID-Capsule before putting the reader on. In each instance, the participants put the reader on <1 minute after simulating ingestion of the ID-Capsule. Because of the approximately 30-minute detection window, these instances would not have resulted in a missed ingestion event. In subsequent scenarios, all participants remembered to wear their reader before taking the ID-Capsule • A participant showed initial difficulty in recognizing the white indicator light; in subsequent scenarios, the participant was able to recognize the white light without difficulty • Another participant had difficulty recognizing the white indicator light and prematurely removed the reader during an ingestion event. This participant showed no later difficulties related to this task for the remainder of the test session • Several participants had difficulty initially interpreting a blinking versus steady reader indicator light but had no further difficulty in the test session • A participant wore the reader incorrectly with the gold side of the reader facing away from the body based on instruction that they thought they had received from the training videos. After referencing the quick start guide, the participant self-corrected. The training videos were reviewed, and there was only mention that the reader should be worn with the gold side facing the body
PS03	Record ingestion event using the ID-Cap reader: Co-ingested ID-Capsule with medication	100 (17/17)	• Some participants placed the reader on the charging pad in the wrong orientation but self-corrected and used the correct orientation for the remainder of the test session
PS04	App & reader setup	100 (17/17)	• Two participants initially had difficulty pairing the reader to the ID-Cap App because they had not turned on the reader but self-corrected after consulting the quick start guide. They were able to complete the scenario successfully • A participant had difficulty understanding the iOS Bluetooth pairing request message displayed on their iPhone but was eventually able to pair the reader to the ID-Cap App successfully
PS05	Record ingestion event using the ID-Cap Reader & App: ID-Capsule alone	100 (17/17)	• A participant showed difficulty in understanding the purpose and appropriateness of recording a manual ingestion event, but there is no associated safety risk with this action
PS06	Charge reader	100 (17/17)	• A participant had difficulty distinguishing the blinking orange reader indicator light because of possible poor vision and possible expectancy bias because they stated that they had expected the video to show a green blinking light. The participant correctly answered the appropriate action to take if the reader light is blinking orange
PK[b]01	Interpreting key indicator light	82 (14/17)	• Four participants incorrectly stated that they would place the reader on the charging pad when the reader indicator light was red. After being directed to the quick start guide, each participant was able to understand the meaning of the red light and stated that they would leave the reader off the charging pad

Scenario	Title	Successful completion rate, % (number of participants who successfully completed the task/number of participants who attempted the task)	Summary of use errors, close calls, or use difficulties
PS07	Record ingestion event using the ID-Cap Reader & App: Co-ingested ID-Capsule with medication	100 (17/17)	• The same participant as in PS05 again showed difficulty in understanding the purpose and appropriateness of recording a manual ingestion event, but there is no associated safety risk with this action
PS08	View and interpret ingestion history in app	100 (17/17)	• Two participants had difficulty interpreting the meaning of the icon used to represent a manually recorded ingestion event. After referencing the user guide, both participants were able to find, and understand the meaning of, the icon
PK02	Respond to ID-Cap App reminders – ID-Capsule alone	94 (16/17)	• A participant stated that if they could not remember whether they had taken a once-daily prescribed ID-Capsule, they would take an additional ID-Capsule because they were confident about manually entering the information if the reader did not record the event and because there was no possibility that they would forget. In the following scenario, where the ID-Capsule was taken with a medication, they indicated that they would NOT take a second one and instead contact their physician
PK03	Respond to ID-Cap App reminders – Co-ingested ID-Capsule with medication	100 (16/16)[c]	• None
PK04	Understand key labeling related to the system	100 (15/15)[c]	• None

[a]PS: patient scenario.

[b]PK: patient knowledge.

[c]Time constraints prevented 2 participants from completing use scenarios PK03 or PK04.

Participant Feedback

The participants grasped the potential value of the system in helping them track and report medication adherence and could see the benefits to people who take medications regularly (including themselves). Many commented on the simplicity and ease of use of the system and liked the training and user resources that were provided, especially the training videos and quick start guides.

The participants, in general, conceptually understood that they should always take their medication as prescribed, supporting the intended role of the system as an adjunct tracking system that does not replace or change physician instructions. Of the 16 participants who answered the postsession interview questions (1 participant did not complete the postsession interview), 16 (100%) stated that they felt that they could use the system effectively and safely to record and track ingestion events and 16 (100%) reported that they believed that the system is safe to use *as is* (Textbox 1).

Textbox 1. Selected patient verbatim quotes reflective of overall participant feedback.

Relevant quotes from patient users

- "I think it's a really great idea for people who have issues with remembering medications and for MDs tracking how they're doing with taking those medications." [P03]
- "Pretty slick! Pretty minimalist, which is good for the target audience. Very easy. One button [on reader] is good. Easy to use, easy to set up. Not a lot to do. Seems very user friendly." [P09]
- "It's a tool. For people that have to take a lot of meds, it is a good tool. I can see the value if it communicates to a provider." [P15]
- "It's simple to use. If everything's working correctly, it meets its intended purpose. The videos and instructions are very good. I had no problems trying to figure it out and follow the process." [P05]
- "Having test driven it, I have confidence in it. That equates [to] safe use." [P05]

Discussion

Principal Findings

The patient's voice has become an important one in health care. Patients are increasingly involved in decisions about their health and medical treatment, and they have become sophisticated health care technology users who understand the value of digital platforms and are eager to use them.

Certainly, the value of the ID-Cap System and other remotely deployed digital health solutions depends on the willingness of patients to engage with them and the ability of patient users to effectively, safely, and conveniently incorporate them into everyday life with minimal training and oversight. The US Food and Drug Administration requires human factors validation testing of digital pill systems and many other medical devices to ensure safety and effectiveness for a device's users, uses, and intended use environments [20].

The results of this human factors validation study show that a representative group of patient users successfully completed the critical and safety-related use tasks necessary for optimal use of the system independently, after receiving training that was followed by a period of training decay. Although the participants were representative of a diverse group of potential patient users who regularly take oral prescription medications, the sample size in this study was limited. The participants ranged in age from 27 to 74 years (mean 51 years, SD 13.8 years), and they were heterogeneous in other demographic factors as well, such as education level (3/17, 18% reported high school graduate as highest education level), handedness (3/17, 18% were left-handed), and sex (7/17, 41% were women). This study was neither designed nor powered to evaluate differences in usability based on demographic factors, medical history, or medication use. It is important to note that nearly every screened participant, regardless of age or education level, agreed to participate in this study after receiving information about the digital pill system and successfully completed the use tasks, showing that patient users of all types adapted well to a novel digital pill system. Use errors, when they did occur, occurred only in the first instance for tasks that are repeated with use, indicating that the participants learned to use the ID-Cap System rapidly—a positive prognosticator for real-world use. Most patients expressed satisfaction with the ID-Cap System and responded favorably to questions about the ease of use and the perceived value of the system.

Prior iterations of digital pill systems used a patch-based reader that adhered to the patient's skin. Clinical evaluations of the patch-based form factor indicated significant limitations from the patient user's perspective with respect to tolerability and usability [17,21]. The ID-Cap System that was tested uses a reader on a lanyard. The patient users found this reader to be easy to use and acceptable in its current form. A wrist-worn reader that may be worn like a watch or may be attached or integrated into the user's existing watch or smartwatch is currently being evaluated. Patients are not only adjusting to new therapeutic regimens but also working to develop new medication-taking behaviors and to adopt support tools and programs that will help them to be successful. Readers that are unobtrusive and can be easily incorporated into daily life will be most readily accepted [18].

The limitations of our study include the fact that this was a simulated use of the ID-Cap System and did not include the actual taking of medication for adherence tracking. The participants reported taking prescription medications by mouth on a regular basis. However, they were not asked to use the digital pill system in this study to actually track and record their own medication use in the same way that patient users would use it in the real world.

In most clinical applications, patient users would use the system chronically over extended periods of time. This validation test was limited to only a few simulated ingestion events and did not evaluate use over time. Certainly, there may be specific patient populations, medical conditions, or treatment-related effects that would affect the usability of the ID-Cap System. This study evaluated the general operation of the device across a diverse group of patient users with different health conditions and varied medication history. Use-related risks should be assessed when the device is applied to specific clinical situations and patient populations to ensure continued safe and effective use from a human factors perspective. Digital pill systems may be incorrectly used or misused by patients in clinical trials or in clinical practice; however, these aspects of use and failure modes were not specifically explored in this study. Risk analyses have been conducted by the device manufacturer to examine and mitigate risks to patients and device performance associated with device use and misuse.

The availability of a call center or additional supporting resources during the use of the device may assist patients, as would engagement with their health care provider. These resources were not evaluated in this test protocol. Patient support programs and data-driven interventions offered by the care team members, research personnel, or device manufacturer may be beneficial for the use of the ID-Cap System in clinical practice and clinical research. Continued efforts to educate clinicians and patients alike regarding the value of the information provided by the system and proper use of the system will further enhance its adoption. In addition, integration of the ID-Cap System into existing care models, electronic health records, and clinical data management systems is being explored.

There is great potential for digital pill systems such as the ID-Cap System to contribute to the efficiency and ultimate success of clinical trial programs. For example, digital pills may be used to assess the likelihood of adherent behavior among prospective trial participants. Once a trial has started, digital pills can identify patient nonadherence and changes in patterns of use early, enabling rapid intervention and course correction. Dose-finding studies for self-administered oral medications are another ideal application for digital pills because their results and outcomes are dependent on (1) human behavior as it relates to medication taking and (2) the quality of adherence measurement to optimize drug exposure and dosing decisions. The robustness of the adherence data collected with digital pills provides an added level of reassurance that the efficacy and safety results of pivotal drug development trials eventually reported are accurate. The inconsistency of medication

adherence reporting within clinical trials has led to the creation of guidelines for researchers and trial sponsors on the inclusion and implementation of adherence measures in study protocols [12-14,22].

Conclusions

The extent of medication adherence, both in clinical use and in clinical trials, is a controllable factor important for therapeutic success and drug development. The pursuit of a solution to the widespread problem of medication nonadherence has led to digital pill systems, which have shown strong performance and a high rate of accuracy. Currently, >15 years of experience and safety data support the use of digital pills with >140,000 ingestions recorded in >1000 patients who have used these devices safely and effectively [23]. In this human factors validation study, the patient users demonstrated the ability to rapidly learn how to use the ID-Cap System and to safely and effectively use the system as intended. The patient users concluded that the device was easy to use and had the potential to be a useful tool for helping to manage their medications. As health care continues to evolve toward remote care delivery and digital health solutions become ubiquitous, systems such as the ID-Cap System that are easy to use, accepted by patients, and valuable in achieving health outcomes will be indispensable.

Acknowledgments

This usability validation test was sponsored and funded by etectRx, Inc. Tensentric, Inc, conducted the validation test at its facility in Boulder, Colorado, and, in collaboration with etectRx, reported the test results. A third-party recruiting firm in Boulder, Colorado, was responsible for participant recruitment and screening. This manuscript was developed by etectRx, Inc. The authors wish to thank Ms Kate Jones of Global MedCom Consulting, LLC, for her editorial assistance.

Conflicts of Interest

SB, CC, and EB are employees of etectRx, Inc. ER, JJ, and EW are employees of Tensentric, Inc. KJ is a consultant for etectRx, Inc.

References

1. Agot K, Taylor D, Corneli AL, Wang M, Ambia J, Kashuba AD, et al. Accuracy of self-report and pill-count measures of adherence in the fem-prep clinical trial: implications for future hiv-prevention trials. AIDS Behav 2015 May;19(5):743-751 [FREE Full text] [doi: 10.1007/s10461-014-0859-z] [Medline: 25100053]
2. Kekäle M, Talvensaari K, Koskenvesa P, Porkka K, Airaksinen M. Chronic myeloid leukemia patients' adherence to peroral tyrosine kinase inhibitors compared with adherence as estimated by their physicians. Patient Prefer Adherence 2014;8:1619-1627 [FREE Full text] [doi: 10.2147/PPA.S70712] [Medline: 25473270]
3. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 4;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]
4. Lloyd JT, Maresh S, Powers CA, Shrank WH, Alley DE. How much does medication nonadherence cost the Medicare fee-for-service program? Med Care 2019 Mar;57(3):218-224. [doi: 10.1097/MLR.0000000000001067] [Medline: 30676355]
5. Pinsky BW, Takemoto SK, Lentine KL, Burroughs TE, Schnitzler MA, Salvalaggio PR. Transplant outcomes and economic costs associated with patient noncompliance to immunosuppression. Am J Transplant 2009 Nov;9(11):2597-2606 [FREE Full text] [doi: 10.1111/j.1600-6143.2009.02798.x] [Medline: 19843035]
6. Makubate B, Donnan PT, Dewar JA, Thompson AM, McCowan C. Cohort study of adherence to adjuvant endocrine therapy, breast cancer recurrence and mortality. Br J Cancer 2013 Apr 16;108(7):1515-1524 [FREE Full text] [doi: 10.1038/bjc.2013.116] [Medline: 23519057]
7. Greer JA, Amoyal N, Nisotel L, Fishbein JN, MacDonald J, Stagl J, et al. A systematic review of adherence to oral antineoplastic therapies. Oncologist 2016 Mar;21(3):354-376 [FREE Full text] [doi: 10.1634/theoncologist.2015-0405] [Medline: 26921292]
8. de Mendoza AH, Cabling ML, Dilawari A, Turner JW, Fernández N, Henderson A, et al. Providers' perspectives on adherence to hormonal therapy in breast cancer survivors. Is there a role for the digital health feedback system? Health Technol (Berl) 2019 Mar;9(2):175-184 [FREE Full text] [doi: 10.1007/s12553-018-0267-x] [Medline: 31448187]
9. Shiovitz TM, Bain EE, McCann DJ, Skolnick P, Laughren T, Hanina A, et al. Mitigating the effects of nonadherence in clinical trials. J Clin Pharmacol 2016 Sep;56(9):1151-1164 [FREE Full text] [doi: 10.1002/jcph.689] [Medline: 26634893]
10. 8 reasons patients don't take their medications. American Medical Association. 2020. URL: https://www.ama-assn.org/delivering-care/patient-support-advocacy/8-reasons-patients-dont-take-their-medications [accessed 2021-04-26]
11. Polinski JM, Kesselheim AS, Frolkis JP, Wescott P, Allen-Coleman C, Fischer MA. A matter of trust: patient barriers to primary medication adherence. Health Educ Res 2014 Oct;29(5):755-763. [doi: 10.1093/her/cyu023] [Medline: 24838119]
12. Vrijens B, Urquhart J. Methods for measuring, enhancing, and accounting for medication adherence in clinical trials. Clin Pharmacol Ther 2014 Jun;95(6):617-626. [doi: 10.1038/clpt.2014.59] [Medline: 24739446]
13. Murali KM, Mullan J, Chen JH, Roodenrys S, Lonergan M. Medication adherence in randomized controlled trials evaluating cardiovascular or mortality outcomes in dialysis patients: a systematic review. BMC Nephrol 2017 Jan 31;18(1):42 [FREE Full text] [doi: 10.1186/s12882-017-0449-1] [Medline: 28143438]

14. Blaschke TF, Osterberg L, Vrijens B, Urquhart J. Adherence to medications: insights arising from studies on the unreliable link between prescribed and actual drug dosing histories. Annu Rev Pharmacol Toxicol 2012;52:275-301. [doi: 10.1146/annurev-pharmtox-011711-113247] [Medline: 21942628]
15. Eliasson L, Clifford S, Mulick A, Jackson C, Vrijens B. How the EMERGE guideline on medication adherence can improve the quality of clinical trials. Br J Clin Pharmacol 2020 Apr;86(4):687-697 [FREE Full text] [doi: 10.1111/bcp.14240] [Medline: 32034923]
16. Aldeer M, Martin R. Medication adherence monitoring using modern technology. In: Proceedings of the IEEE 8th Annual Ubiquitous Computing, ElectronicsMobile Communication Conference (UEMCON). 2017 Oct Presented at: IEEE 8th Annual Ubiquitous Computing, ElectronicsMobile Communication Conference (UEMCON); Oct. 19-21, 2017; New York, NY, USA p. 491-497. [doi: 10.1109/uemcon.2017.8249101]
17. Eisenberger U, Wüthrich RP, Bock A, Ambühl P, Steiger J, Intondi A, et al. Medication adherence assessment: high accuracy of the new ingestible sensor system in kidney transplants. Transplantation 2013 Aug 15;96(3):245-250 [FREE Full text] [doi: 10.1097/TP.0b013e31829b7571] [Medline: 23823651]
18. Chai P, Castillo-Mancilla J, Buffkin E, Darling C, Rosen R, Horvath K, et al. Utilizing an ingestible biosensor to assess real-time medication adherence. J Med Toxicol 2015;11(4):439-444 [FREE Full text] [doi: 10.1007/s13181-015-0494-8] [Medline: 26245878]
19. Weinger M, Wiklund M, Gardner-Bonneau D. Handbook of Human Factors in Medical Device Design. Boca Raton: CRC Press; 2010.
20. Applying human factors and usability engineering to medical devices: guidance for industry and food and drug administration staff. U.S. Department of Health and Human Services, Food and Drug Administration. 2016. URL: https://www.fda.gov/media/80481/download [accessed 2021-05-07]
21. ABILIFY MYCITE® (aripiprazole tablets with sensor) - manufactured by Otsuka Pharmaceutical Co. Ltd., Tokyo. US Food and Drug Administration. 2020. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/207202s003lbl.pdf [accessed 2021-09-03]
22. Breckenridge A, Aronson JK, Blaschke TF, Hartman D, Peck CC, Vrijens B. Poor medication adherence in clinical trials: consequences and solutions. Nat Rev Drug Discov 2017 Mar;16(3):149-150. [doi: 10.1038/nrd.2017.1] [Medline: 28154411]
23. Plowman RS, Peters-Strickland T, Savage GM. Digital medicines: clinical review on the safety of tablets with sensors. Expert Opin Drug Saf 2018 Sep;17(9):849-852. [doi: 10.1080/14740338.2018.1508447] [Medline: 30073875]

Edited by A Kushniruk; submitted 28.05.21; peer-reviewed by C Pathiravasan , D McCann, H Mehdizadeh; comments to author 04.07.21; revised version received 07.07.21; accepted 10.07.21; published 08.11.21.

Please cite as:
Baumgartner SL, Buffkin Jr DE, Rukavina E, Jones J, Weiler E, Carnes TC
A Novel Digital Pill System for Medication Adherence Measurement and Reporting: Usability Validation Study
JMIR Hum Factors 2021;8(4):e30786
URL: https://humanfactors.jmir.org/2021/4/e30786
doi:10.2196/30786
PMID:34747709

Original Paper

The Effectiveness of a Multidisciplinary Electronic Discharge Readiness Tool: Prospective, Single-Center, Pre-Post Study

Angela Keniston[1*], MSPH; Lauren McBeth[1*], BA; Jonathan Pell[1*], MD; Kasey Bowden[1*], MSN, FNP, AG-ACNP; Anna Metzger[2*], MPH; Jamie Nordhagen[3], MS, RN; Amanda Anthony[3*], DNP, CCNS, ACNS-BC; John Rice[4*], PhD; Marisha Burden[1*], MD

[1]Anschutz Medical Campus, Division of Hospital Medicine, University of Colorado, Aurora, CO, United States

[2]Colorado School of Public Health, University of Colorado, Aurora, CO, United States

[3]UCHealth, Denver, CO, United States

[4]Adult and Child Center for Health Outcomes Research and Delivery Science, University of Colorado, Aurora, CO, United States

[*]these authors contributed equally

Corresponding Author:
Angela Keniston, MSPH
Anschutz Medical Campus
Division of Hospital Medicine
University of Colorado
12401 E. 17th Avenue
Mail Stop F782
Aurora, CO, 80045
United States
Phone: 1 7202401431
Email: Angela.Keniston@cuanschutz.edu

Abstract

Background: In the face of hospital capacity strain, hospitals have developed multifaceted plans to try to improve patient flow. Many of these initiatives have focused on the timing of discharges and on lowering lengths of stay, and they have met with variable success. We deployed a novel tool in the electronic health record to enhance discharge communication.

Objective: The aim of this study is to evaluate the effectiveness of a discharge communication tool.

Methods: This was a prospective, single-center, pre-post study. Hospitalist physicians and advanced practice providers (APPs) used the Discharge Today Tool to update patient discharge readiness every morning and at any time the patient status changed throughout the day. Primary outcomes were tool use, time of day the clinician entered the discharge order, time of day the patient left the hospital, and hospital length of stay. We used linear mixed modeling and generalized linear mixed modeling, with team and discharging provider included in all the models to account for patients cared for by the same team and the same provider.

Results: During the pilot implementation period from March 5, 2019, to July 31, 2019, a total of 4707 patients were discharged (compared with 4558 patients discharged during the preimplementation period). A total of 352 clinical staff had used the tool, and 84.85% (3994/4707) of the patients during the pilot period had a discharge status assigned at least once. In a survey, most respondents reported that the tool was helpful (32/34, 94% of clinical staff) and either saved time or did not add additional time to their workflow (21/24, 88% of providers, and 34/34, 100% of clinical staff). Although improvements were not observed in either unadjusted or adjusted analyses, after including starting morning census per team as an effect modifier, there was a reduction in the time of day the discharge order was entered into the electronic health record by the discharging physician and in the time of day the patient left the hospital (decrease of 2.9 minutes per additional patient, P=.07, and 3 minutes per additional patient, P=.07, respectively). As an effect modifier, for teams that included an APP, there was a significant reduction in the time of day the patient left the hospital beyond the reduction seen for teams without an APP (decrease of 19.1 minutes per patient, P=.04). Finally, in the adjusted analysis, hospital length of stay decreased by an average of 3.7% (P=.06).

Conclusions: The Discharge Today tool allows for real time documentation and sharing of discharge status. Our results suggest an overall positive response by care team members and that the tool may be useful for improving discharge time and length of stay if a team is staffed with an APP or in higher-census situations.

(JMIR Hum Factors 2021;8(4):e27568) doi:10.2196/27568

KEYWORDS

discharge planning; health information technology; quasi-experimental design; multidisciplinary; teamwork

Introduction

Hospitals around the country, in particular tertiary and quaternary referral centers, can face bottlenecks and capacity issues [1-3]. Successful management of capacity and throughput by hospitals allows increased access for patients who need a higher level of care and expertise [1-4]. Delayed discharge of hospitalized patients can impede the flow of patients throughout a hospital [1,3,5-8], resulting in delays in care for patients being admitted [9,10] and adverse events, including medication errors [11,12], infections [13], and increased mortality [13-16]. Delays in discharge are associated with both increased lengths of stay and costs [2,17-19].

The commonly used discharge communication workflows often hinder efficient, timely discharge [20]. Many hospitals document an expected date of discharge at the time of admission, and triaging of work is based on this information documented very early in the patient admission process; however, patient condition changes frequently throughout hospitalization [21]. Clinicians, nurses, care management, pharmacy, and other team members often meet midmorning or in the afternoon each day to discuss discharge needs for hospitalized patients; however, minimal communication occurs before these meetings or in real time; in addition, these meetings do not integrate well into workflows [22]. The lack of communication early in the day, before rounding on patients, delays discharge communication and, ultimately, patient discharge. Earlier discharge, by as little as 1 hour, has been shown to alleviate hospital crowding, reduce access blocking, and improve patient flow [23,24].

Typical workflows rely on processes implemented outside of the electronic health record (EHR), such as meetings, paging, and telephone calls, which are inadequate for efficient discharge communication and frequently interrupt patient care [25,26]. Health information technology solutions most often described in the literature include passive communication tools, such as electronic patient journey boards, hospital capacity dashboards, asynchronous electronic reports, and discharge checklists [4,27-34], or health information technology tools that reside outside of the EHR [35,36]. Even commonly used tools within the EHR, such as messaging or conditional discharge orders, do not provide real time, integrated communications despite being a function of an EHR [37,38].

To address these deficits, we developed a novel EHR tool to facilitate communication in real time between hospitalists and other clinicians and care team members about discharge readiness and barriers to discharge. We evaluated whether the use of this tool was associated with improvements in discharge order time, discharge time, and length of stay. In addition, we evaluated whether this tool worked differently under different conditions, such as high-census days or when an advanced practice provider (APP) was assigned to a patient team. Finally, we evaluated whether the effects of this tool persisted after formal stakeholder engagement efforts waned.

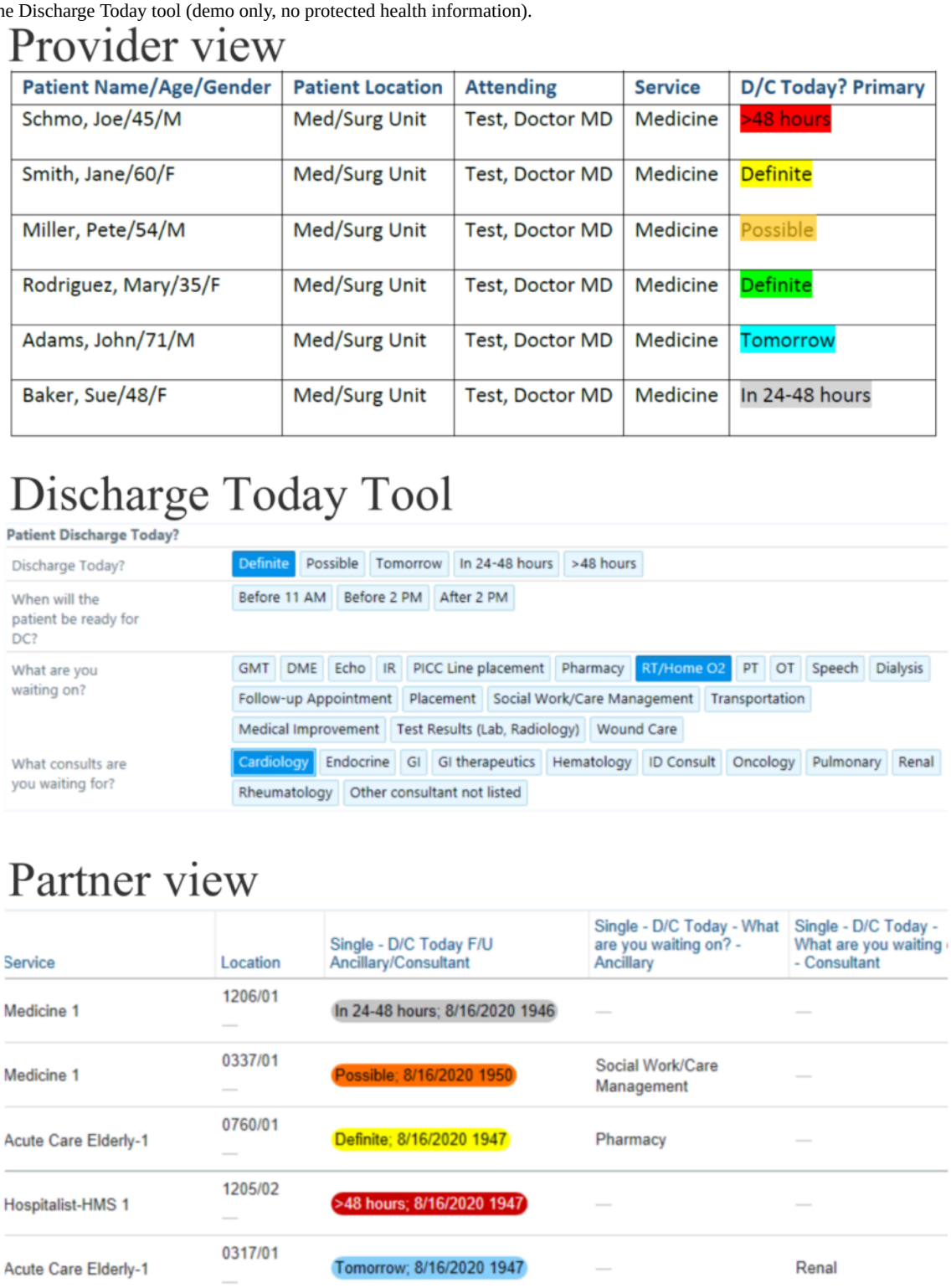
Methods

Tool Development

Using multiple user-centered design strategies [39-42], the Discharge Today Tool was iteratively developed from July 1, 2018, to July 31, 2019, and deployed to hospitalists and other clinical staff on March 5, 2019. This tool was designed to integrate with customizable EHR patient worklists used by most clinicians and staff members providing clinical care to hospitalized patients (Figure 1).

In the provider view, hospitalists may access the Discharge Today tool via the D/C Today? Primary column in their EHR patient list. Using this tool, hospitalists may document patient discharge readiness (definite today, possible today, tomorrow, in 24 to 48 hours, or in more than 48 hours) and if the hospitalist is waiting on any final care before the patient can be discharged. Via the partner view, the data collected by the Discharge Today tool is shared with ancillary and consulting clinicians in the Single—D/C Today—What are you waiting on?—Ancillary and the Single—D/C Today—What are you waiting on?—Consultant columns in their EHR patient worklists. The definitions for the discharge readiness statuses are as follows: Definite-very high probability that the patient will be discharged today unless there are unexpected changes during the day. For example, if you have a patient who is clinically ready for discharge but needs home oxygen set up, this patient would be considered a definite discharge, awaiting respiratory therapy. Possible-some probability that the patient could be discharged today. For example, if you have a patient with complex health conditions waiting for subacute nursing facility placement, this patient would be considered a possible discharge, awaiting placement. Tomorrow: very likely that the patient could be discharged tomorrow. In 24-48 hours: the patient is not going home today but will likely be discharged in the next 24 to 48 hours; >48 hours: very unlikely that the patient would be discharged within the next 48 hours.

Figure 1. The Discharge Today tool (demo only, no protected health information).



Provider view

Patient Name/Age/Gender	Patient Location	Attending	Service	D/C Today? Primary
Schmo, Joe/45/M	Med/Surg Unit	Test, Doctor MD	Medicine	>48 hours
Smith, Jane/60/F	Med/Surg Unit	Test, Doctor MD	Medicine	Definite
Miller, Pete/54/M	Med/Surg Unit	Test, Doctor MD	Medicine	Possible
Rodriguez, Mary/35/F	Med/Surg Unit	Test, Doctor MD	Medicine	Definite
Adams, John/71/M	Med/Surg Unit	Test, Doctor MD	Medicine	Tomorrow
Baker, Sue/48/F	Med/Surg Unit	Test, Doctor MD	Medicine	In 24-48 hours

Discharge Today Tool

Patient Discharge Today?

Discharge Today?	Definite; Possible; Tomorrow; In 24-48 hours; >48 hours
When will the patient be ready for DC?	Before 11 AM; Before 2 PM; After 2 PM
What are you waiting on?	GMT; DME; Echo; IR; PICC Line placement; Pharmacy; RT/Home O2; PT; OT; Speech; Dialysis; Follow-up Appointment; Placement; Social Work/Care Management; Transportation; Medical Improvement; Test Results (Lab, Radiology); Wound Care
What consults are you waiting for?	Cardiology; Endocrine; GI; GI therapeutics; Hematology; ID Consult; Oncology; Pulmonary; Renal; Rheumatology; Other consultant not listed

Partner view

Service	Location	Single - D/C Today F/U Ancillary/Consultant	Single - D/C Today - What are you waiting on? - Ancillary	Single - D/C Today - What are you waiting - Consultant
Medicine 1	1206/01	In 24-48 hours; 8/16/2020 1946	—	—
Medicine 1	0337/01	Possible; 8/16/2020 1950	Social Work/Care Management	—
Acute Care Elderly-1	0760/01	Definite; 8/16/2020 1947	Pharmacy	—
Hospitalist-HMS 1	1205/02	>48 hours; 8/16/2020 1947	—	—
Acute Care Elderly-1	0317/01	Tomorrow; 8/16/2020 1947	—	Renal

Hospitalists caring for the patients are able to easily document discharge readiness (definite today, possible today, tomorrow, in 24-48 hours, or in >48 hours) [21] and whether the hospitalist is waiting on any final care before the patient can be discharged. The data collected by the Discharge Today tool are also disseminated via EHR patient worklists, which are EHR-based reports designed to summarize patient care for clinicians using the EHR, and via an automatic paging functionality directly from the EHR. Details can be found in our study describing stakeholder engagement and the user-centered design approaches applied [43].

Addressing communication challenges by improving the efficiency and accuracy of communication may reduce inefficiencies and errors in health care, including during the discharge process. The Discharge Today tool fosters flexibility and agility in communication, including asynchronous communication, feedback loop capabilities, different

functionalities according to user role, and allowing for both formal and informal communication.

Study Design

This study was conducted as a prospective, single-center, quasi-experimental, pre-post study designed to evaluate the effectiveness of the Discharge Today tool. The study was approved by the Colorado Multiple institutional review board as a quality improvement project and funded by a small pilot grant.

Setting

This study was conducted at the University of Colorado Hospital, a 678-bed tertiary care center with approximately 12,000 medicine discharges per year.

Inclusion and Exclusion Criteria

Hospitalist physicians and APPs were trained as they started on service and asked to use the Discharge Today tool every day that they were on service with all patients assigned to their team. Clinicians were asked to update patient discharge readiness statuses first thing in the morning and throughout the day as discharge readiness and needs evolved. Patients who were expected to be discharged >48 hours out only needed an update every 3 days as the tool would automatically unpopulate the patient status if unchanged after 3 days to ensure the most accurate and up-to-date information. Clinicians received a small incentive for participation (ie, coffee or other small tokens of gratitude that were funded by the small grant).

Patients were enrolled in this study as part of their regular hospitalization if they presented during the study period. Patients already in the hospital at the start of the pilot implementation period (March 5, 2019, to July 31, 2019) were excluded from the analysis. Patients admitted on or after March 5, 2019, and discharged on or before July 31, 2019, were assigned to the pilot implementation period.

Data Collection

All patient-level clinical and quality outcomes data queried from the hospital EHR data warehouse were collected as part of their hospitalization process. We queried data from the EHR data warehouse for any patient admitted to the hospital and assigned to a hospital medicine service during the preimplementation period (October 1, 2018, to March 4, 2019), the pilot implementation period (March 5, 2019, to July 31, 2019), and the postimplementation maintenance period (August 1, 2019, to December 31, 2019).

To assess adoption, we documented the number of users who added the tool to their patient worklists within the EHR. To assess both reach and implementation, we queried each time data were entered into the tool by a clinician, including discharge readiness status, when patients assigned a definite discharge status would be ready to be discharged, and what ancillary services or tasks might be needed, such as rehabilitation services, respiratory therapy, pharmacy, social work, care management, medical improvement, or consultant services (Table 1).

Table 1. Definitions.

Variable	Definition	Type	Level
Discharge order time	The time of day the physician entered a discharge order for a patient into the electronic health record	Outcome	Patient encounter
Discharge time	The time of day the patient left the hospital after being discharged	Outcome	Patient encounter
Length of stay	The duration, in hours, between admission to the hospital and discharge from the hospital	Outcome	Patient encounter
Team assignment	The team to which the patient was assigned when they were discharged from the hospital	Random effect	Team
Physician	The physician who discharged the patient	Random effect	Physician
Type of patient	Patients admitted for inpatient hospitalization or patients admitted for observation	Confounder	Patient encounter
Charlson Comorbidity Index	A measure of patient acuity based on patient age and discharge diagnosis ICD-10[a] codes assigned after discharge	Confounder	Patient encounter
Discharge to postacute care	Discharge to a setting other than home, including skilled nursing facilities, hospice, and long-term care	Confounder	Patient encounter
Teaching service	Teams that are staffed with a medical student or resident	Confounder	Team
Staffed with an APP[b]	Teams that are staffed with a physician and an APP	Confounder	Team
Starting morning census	The number of patients assigned to a team at 7 AM each morning	Confounder	Team

[a]ICD-10: International Classification of Diseases, 10th Revision.

[b]APP: advanced practice provider.

Surveys were conducted using REDCap (Research Electronic Data Capture)—a secure, web-based application for building and managing web-based surveys and databases [44]—to evaluate the usability of and experience with hospital medicine physicians, APPs, nurses, care management, and other clinical staff during the pilot implementation period. The complete survey results are reported in a study describing the stakeholder

engagement and user-centered design approaches that we applied [43].

Outcomes

Primary outcomes for assessing the effectiveness of this tool were (1) time of day the physician entered the discharge order, (2) time of day the patient left the hospital, and (3) hospital length of stay. Secondary outcomes were (1) proportion of patients for whom a discharge order was entered before 11 AM and (2) proportion of patients discharged before 11 AM, both metrics commonly used to evaluate patient flow. We also queried our data warehouse for the type of patient (inpatient or observation patient), Charlson Comorbidity Index, type of team (physician alone, physician with APP, physician with resident, or physician with APP and resident), proportion of days in the hospital that discharge status was documented for each patient (0%-25%, 25%-50%, 50%-75%, or >75%), and the number of patients assigned to a team at 7 AM (starting morning census).

Study Size

On the basis of the original planned interrupted time series design, to maximize feasibility against sample size, we allowed for approximately 20 weeks of data collection during each period; that is patients discharged during the preintervention period, patients discharged during the pilot intervention period, and patients discharged during the postintervention period. On the basis of data from 2017, we anticipated an average of approximately 140 discharges per week. However, to account for clustering within providers and teams, the analysis shifted to a mixed modeling approach. Although no post hoc power analysis was conducted, >4000 patients were discharged in each time period.

Data Analysis

We estimated means and SDs for continuous variables when approximately normally distributed (as assessed by visual inspection of histograms), medians and IQRs when not, and frequencies for categorical variables. Descriptive statistics were computed for patient, clinician, and team characteristics.

Patient-level, clinician-level, and team-level covariates, hypothesized a priori to be associated with the time of discharge order, time of discharge, and hospital length of stay, were included in multivariable analyses. Models for discharge order time, actual discharge time, and hospital length of stay were adjusted for (1) type of patient, (2) Charlson Comorbidity Index, (3) teaching service, (4) staffed with an APP, (5) discharge to postacute care, (6) starting morning census per team, (7) team, and (8) physician (Table 1). The discharge order time and discharge time models were also adjusted for hospital length of stay.

We used linear mixed modeling for the analysis of the time of day the hospitalist physician entered the discharge order into the EHR, the time of day the patient left the hospital, and the hospital length of stay. We converted time to hours elapsed since midnight on a 24-hour clock for modeling. For our binary outcomes, specifically, whether a discharge order was entered before 11 AM and whether a patient was discharged before 11 AM, a generalized linear mixed model with logit link function and binary response distribution was used. The intervention period, that is preimplementation and pilot implementation, was the independent variable of interest. Team and discharging physicians were included as random effects in all models to account for correlation between patients cared for by the same team and the same physician. Given that hospital length of stay is right skewed, this variable was log-transformed to facilitate regression analysis. We reported a relative difference in hospital length of stay by exponentiating the coefficient, subtracting 1, and expressing the result as a percentage [45].

Secondary analyses were performed to determine whether potential effect modification was supported by the data. We hypothesized that the Discharge Today tool would help hospitalist physicians with a high number of patients on their team triage work and enter discharge orders more quickly. To test this hypothesis, we included an interaction term between the team starting morning census and intervention period, allowing for the intervention's effect to depend on daily patient volume [21,46]. We also hypothesized that the Discharge Today tool might be more effective for teams staffed with an APP, allowing teams to triage and divide work more efficiently [47-49]. To test this hypothesis, we included an interaction term between whether a team was staffed with an APP and intervention period, allowing the intervention's effect to depend on the presence of an APP.

Patients with missing data on any variables necessary for a specific analysis were excluded from that analysis. All statistical analyses were performed using SAS Enterprise Guide 8.1 (SAS Institute Inc).

Results

Use of the Discharge Today Tool

During the preimplementation period—October 1, 2018, to March 4, 2019—4558 patients were discharged from 1 of 18 hospital medicine teams at the University of Colorado Hospital by 57 hospitalist physicians (Table 2). During the pilot implementation period—March 5, 2019, to July 31, 2019—4707 patients were discharged from 1 of 18 teams by 62 hospitalist physicians.

During the implementation period, 84.85% of the patients discharged were assigned a discharge status. The most common barriers identified were medical improvement, placement, subspecialty consults, physical therapy, and social work or care management (Table 3).

Table 2. Characteristics of teams, clinicians, and patients by project period.

Characteristics	Preimplementation (N=4558)	Pilot implementation (N=4707)
Team type, n (%)[a]		
With APP[b]	2031 (44.56)	2046 (43.47)
Without APP	2527 (55.44)	2661 (56.53)
Teaching	2689 (59.00)	2724 (57.87)
Nonteaching	1869 (41.00)	1983 (42.13)
Discharges per team, mean (SD)	239.9 (115.8)	247.7 (118.7)
Morning census per team, mean (SD)	10.6 (2.6)	10.7 (2.5)
Unique physicians, n (%)	57 (1.25)	62 (1.32)
Discharges per physician, mean (SD)	72.2 (43.7)	69.1 (46.9)
Patient type, n (%)		
Inpatient	3532 (77.49)	3764 (79.97)
Observation patient	1004 (22.03)	919 (19.52)
Missing	22 (0.48)	24 (0.51)
Discharge disposition, n (%)		
Home	3927 (86.16)	4060 (86.25)
Postacute care setting	557 (12.22)	583 (12.39)
Other	62 (1.36)	56 (1.19)
In-hospital death	12 (0.26)	8 (0.17)
Charlson Comorbidity Index, median (IQR)	2 (1-3)	2 (1-3)
Proportion of days in the hospital a Discharge Today tool status was documented for each patient, n (%)		
0%-25% of hospital stay	N/A[c]	401 (8.52)
26%-50% of hospital stay	N/A	1051 (22.33)
51%-75% of hospital stay	N/A	798 (16.95)
>75% of hospital stay	N/A	1253 (26.62)
Missing	N/A	1204 (25.58)

[a]Teams may fall into more than one category; therefore, the total is >100%.

[b]APP: advanced practice provider.

[c]N/A: not applicable.

Table 3. Discharge Today tool use.

Characteristics	Pilot implementation, n (%)
Discharging hospital medicine physicians (n=56)	
Used tool ever	46 (82)
Used tool never	10 (18)
Used always	16 (29)
Patients discharged from a hospital medicine service (n =4707)	
Patients ever assigned a discharge status	3994 (84.85)
Ever definite	2087 (52.25)
Ever possible	2209 (55.31)
Ever tomorrow	N/A[a]
Ever in 24-48 hours	1607 (40.24)
Ever >48 hours	2771 (69.38)
Of the patients ever assigned a discharge status, those with barriers identified	2133 (53.41)
Number of barriers identified (n =4059)	
Medical improvement	1812 (44.64)
Placement	532 (13.11)
Subspecialty consults	365 (8.99)
PT[b]	334 (8.23)
Social work or care management	344 (8.48)
OT[c]	158 (3.89)
RT[d] or home oxygen	159 (3.92)
Transportation	78 (1.92)
Test results (laboratory and radiology)	1 (0.02)
Follow-up appointment	69 (1.70)
IR[e]	66 (1.63)
Echo	30 (0.74)
Dialysis	36 (0.89)
GMT[f]	26 (0.64)
Speech	19 (0.47)
PICC[g] line placement	10 (0.25)
Pharmacy	13 (0.32)
DME[h]	7 (0.17)
Wound care	0 (0)
Discharge Today tool users (n=352)	
Registered nurse	71 (20.2)
Resident	67 (19.0)
Physician	56 (15.9)
Physical therapist	31 (8.8)
Physician assistant	27 (7.7)
Medical student	20 (5.7)
Case manager	18 (5.1)

Characteristics	Pilot implementation, n (%)
Nurse practitioner	15 (4.3)
Occupational therapist	12 (3.4)
Social worker	9 (2.6)
Care coordinator	5 (1.4)
Patient resident liaison	4 (1.1)
Fellow	3 (0.9)
Pharmacist	3 (0.9)
Physical therapy student	3 (0.9)
Respiratory therapist	2 (0.6)
Speech or language pathologist	2 (0.6)
Student nurse	1 (0.3)
Clinical nurse specialist	1 (0.3)
Technician	1 (0.3)
Certified nursing assistant	1 (0.3)

[a]N/A: not applicable.

[b]PT: physical therapy.

[c]OT: occupational therapy.

[d]RT: respiratory therapy.

[e]IR: interventional radiology.

[f]GMT: glucose management team

[g]PICC: peripherally inserted central catheter.

[h]DME: durable medical equipment.

Of the 56 hospitalists who discharged a patient during the pilot implementation period, 46 (82%) used the tool for patients assigned to their teams. During the pilot implementation period, 352 users, including physicians, APPs, residents and medical students, nurses, physical and occupational therapists, care managers and social workers, and pharmacists, added the tool to their patient worklists. Of these users, 86% (48/56) of hospitalist physicians and 88% (29/33) of hospitalist APPs added the tool to their EHR patient lists. Physicians, APPs, residents, and medical students added the primary column in which they entered a discharge readiness status daily, and other clinical staff, including nurses, physical and occupational therapists, care managers and social workers, and pharmacists, added the read-only columns where the discharge readiness status entered by providers can be viewed. In addition, in some cases, the tool was added to shared patient worklists, which meant that >352 clinical staff were using the tool.

Hospital medicine physicians, APPs, nurses, care management, and other clinical staff reported in a survey conducted during the pilot implementation period that the tool did not adversely affect their workflow (21/24, 88% of the providers, and 34/34, 100% of clinical staff) and was helpful for managing the patient discharge process (32/34, 94% of clinical staff).

Effectiveness of the Discharge Today Tool

In both unadjusted effectiveness analysis and after adjusting for prespecified confounders, we did not find a significant reduction in the time of day the discharge order was entered into the EHR by the discharging physician during the pilot implementation period compared with the preimplementation period (Table 4).

In the secondary analyses for effect modification, we observed an interaction effect between intervention period and starting morning census (P=.07; Figure 2).

The time of day the discharge order was entered into the EHR by the discharging physician varied according to the number of patients assigned to a team at 7 AM each morning in the pilot implementation period compared with the preimplementation period.

Specifically, the time of day the discharge order was entered into the EHR by the discharging physician decreased by an additional 2.9 minutes per patient for every 1-patient increase in morning census during the pilot implementation period compared with each 1-patient increase in morning census during the preimplementation period. However, we did not find any evidence of effect modification for the intervention by the presence of an APP (Table 4).

Although in unadjusted and adjusted analyses the time of day the patients left the hospital for the pilot implementation period compared with the preimplementation period did not change significantly, we found, in secondary analyses conducted to investigate effect modification, that the average time of day the patients left the hospital decreased for every 1-patient increase in morning census for a given team during the preimplementation period compared with the pilot implementation period by 3.0 minutes (P=.07; Figure 3).

Table 4. Discharge Today tool effectiveness modeling by project period.

Characteristics	Discharge order time, mean (SD)[a]	Discharge time, mean (SD)[a]	Length of stay in hours, median (IQR)[b]	Discharge order before 11 AM[c], n (%)	Discharge before 11 AM[c], n (%)
Preimplementation (N=4114)	12:40 (2:38)	14:41 (2:46)	75 (47-138)	1125 (27.35)	382 (9.29)
Pilot implementation (N=4285)	12:45 (2:33)	14:44 (2:43)	76 (46-139)	1103 (25.74)	367 (8.56)
Unadjusted results					
95% CI	5.1 (–3.8 to 14.1)	2.1 (–6.6 to 10.9)	1.9 (–6 to 2.3)	0.94 (0.83 to 1.1)	0.91 (0.77 to 1.1)
P value	.26	.63	.37	.35	.26
Adjusted results					
95% CI	6.8 (–2.2 to 15.8)	4.0 (–5.2 to 13.2)	3.7 (–7.4 to 0.1)	0.92 (0.81 to 1.0)	0.90 (0.76 to 1.1)
P value	.14	.39	.06	.19	.22
Starting morning census					
95% CI	–2.9 (–5.9 to 0.2)	–3.0 (–6.2 to –0.2)	0.3 (–1.7 to 1.1)	1.0 (0.97 to 1.1)	1.1 (0.98 to 1.1)
P value	.07	.07	.66	.66	.16
Staffed with an APP[d]					
95% CI	–9.2 (–27.1 to 8.6)	–19.1 (–37 to –0.9)	4.7 (–11.9 to 3.1)	1.1 (0.84 to 1.4)	1.4 (0.99 to 2.0)
P value	.31	.04	.23	.53	.06

[a]Mean difference (in minutes) calculated.

[b]Mean percentage decrease calculated.

[c]Odds ratio calculated.

[d]APP: advanced practice provider.

Figure 2. Discharge order time: interaction between team starting morning census and intervention period (preimplementation vs pilot implementation).

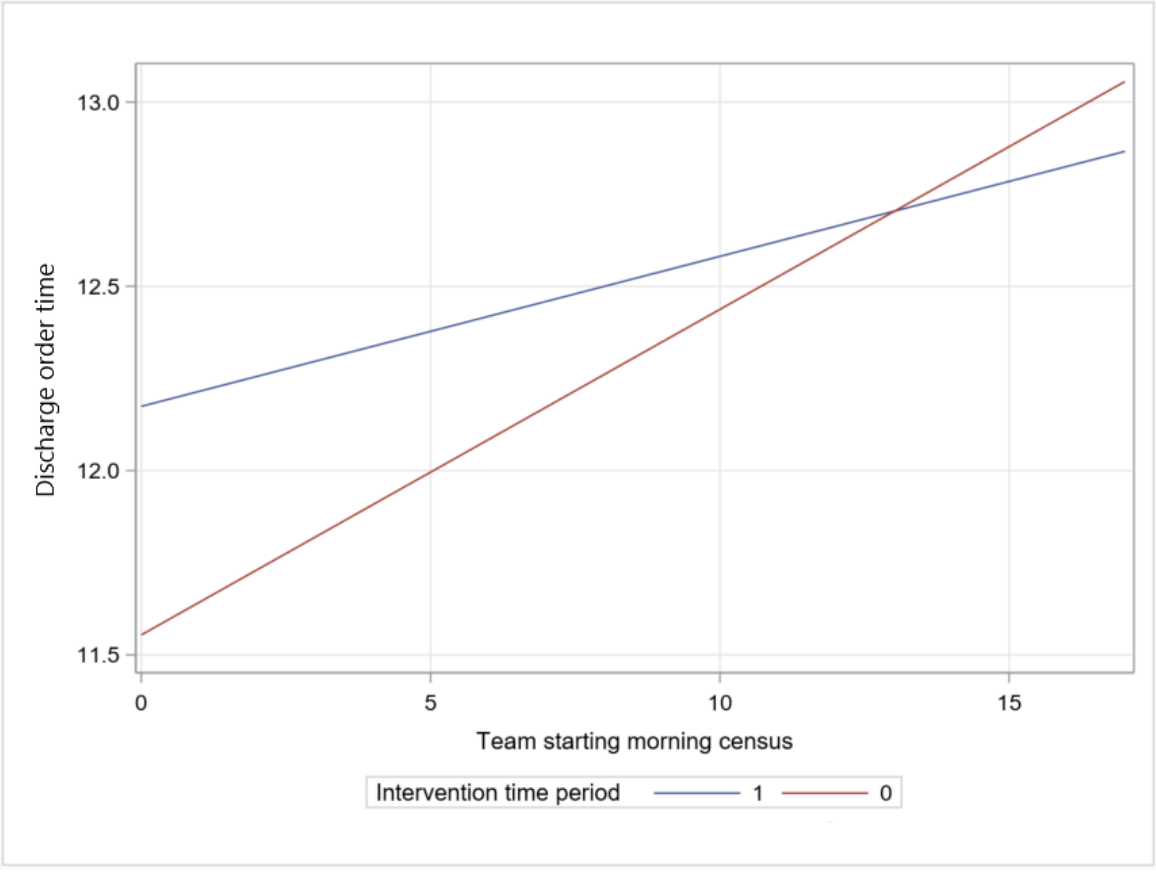
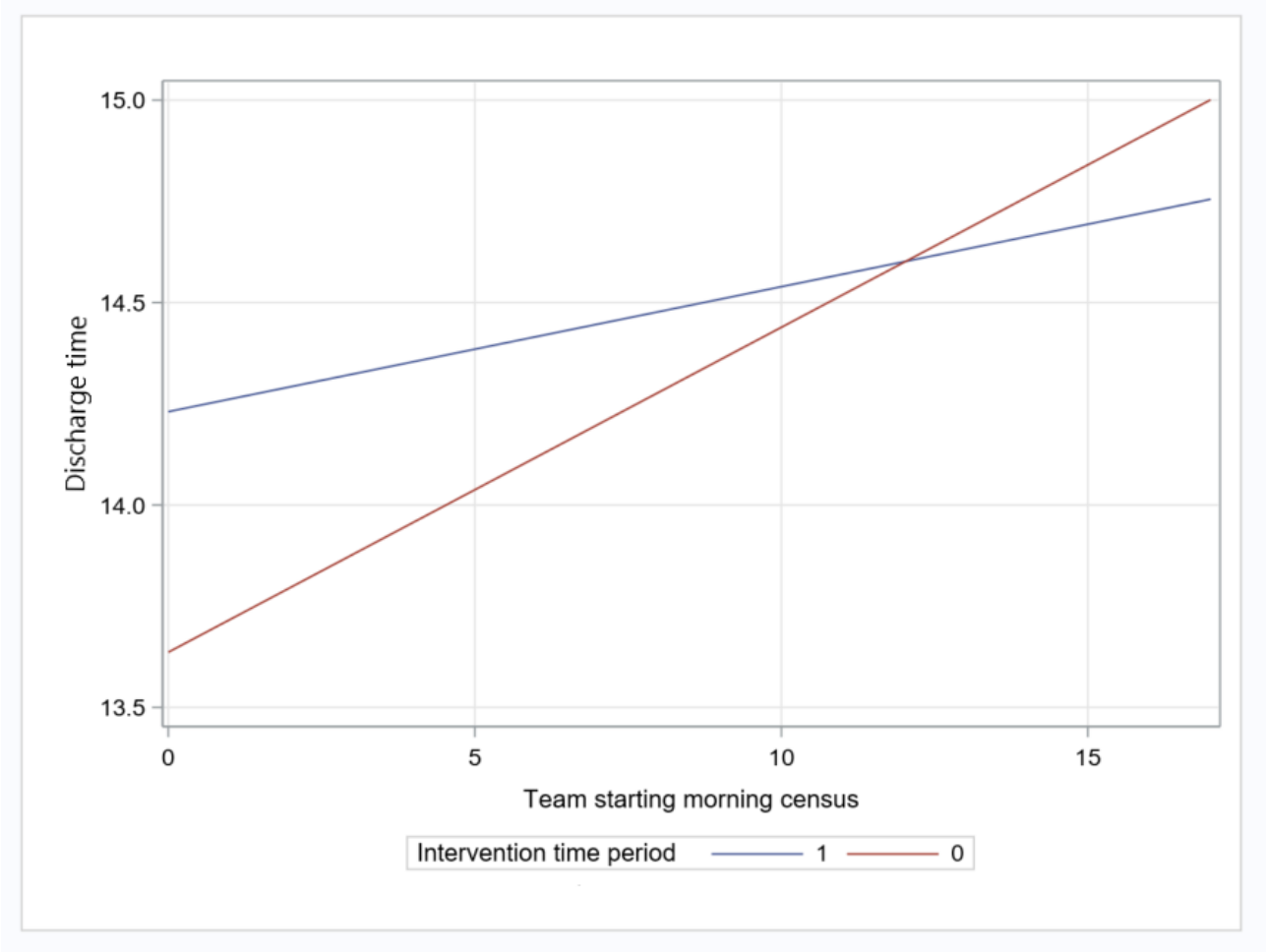


Figure 3. Discharge time: interaction between team starting morning census and intervention period (preimplementation vs pilot implementation).

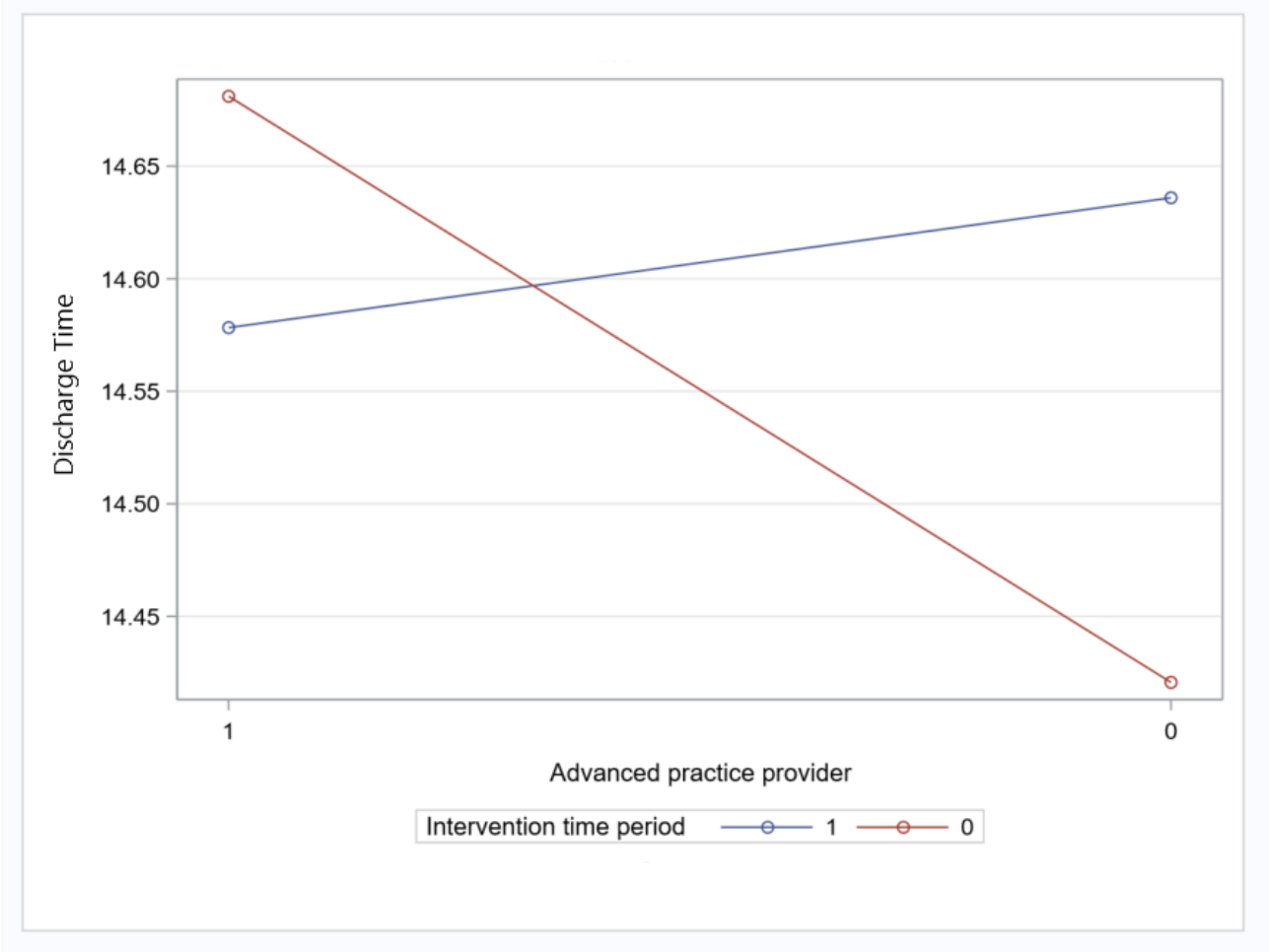


The time of day the patient left the hospital after being discharged varied according to the number of patients assigned to a team at 7 AM each morning in the pilot implementation period compared with the preimplementation period.

In addition, the average time of day the patients left the hospital decreased for teams staffed with an APP during the preimplementation period compared with the pilot implementation period by 19.1 minutes (P=.04; Table 4; Figure 4).

The time of day the patient left the hospital after being discharged varied according to whether a team was staffed with an advanced practice provider in the pilot implementation period compared with the preimplementation period.

Figure 4. Discharge time: interaction between team staffed with an advanced practice provider and intervention period (preimplementation vs pilot implementation).



In the unadjusted analysis, hospital length of stay did not change significantly. After adjusting for prespecified confounders, we observed a trend toward reduction in hospital length of stay for the pilot implementation period compared with the preimplementation period (decrease of 3.7%; P=.06). We did not observe significant changes in the length of stay from preimplementation to pilot implementation under different conditions, such as high-census days or presence of an APP on a patient team; that is, no significant interactions between the intervention period and these variables were detected (Table 4).

Neither of the secondary outcomes—proportion of patients for whom a discharge order was entered before 11 AM and proportion of patients discharged before 11 AM—was found to significantly improve after introduction of the Discharge Today tool in unadjusted analysis, after adjusting for prespecified covariates, or under different conditions (Table 4).

To test whether the effects of this tool persisted in a maintenance period during which stakeholder engagement efforts were curtailed, we compared the outcomes of the pilot implementation period with those of the postimplementation period using mixed effects models (Table 5). Adjusting for prespecified covariates, we observed a significant reduction in the time of day the discharge order was entered into the EHR for teams staffed with an APP during the postimplementation period compared with teams staffed with an APP during the pilot implementation period (an average decrease of 20.1 minutes per patient (95% CI –36.1 minutes to –4.0 minutes; P=.01; Figure 5).

The time of day the discharge order was entered into the EHR by the discharging physician varied according to whether a team was staffed with an advanced practice provider in the postimplementation (maintenance) period compared with the pilot implementation period.

However, no other outcomes improved significantly from the pilot implementation period to the postimplementation period.

Table 5. Discharge Today tool effectiveness modeling comparing pilot implementation and postimplementation periods.

Characteristics	Discharge order time[a], mean (SD)	Discharge time[a], mean (SD)	Length of stay in hours[b], median (IQR)	Discharge order before 11 AM[c], n (%)	Discharge before 11 AM[c], n (%)
Pilot implementation (N=4285)	12:45 (2:33)	14:44 (2:43)	76 (46-139)	1103 (25.74)	367 (8.56)
Postimplementation (N=4255)	12:56 (2:29)	14:53 (2:38)	79 (47-142)	924 (21.72)	327 (7.69)
Unadjusted results					
95% CI	11 (3.1 to 18.9)	9.8 (2.0 to 17.6)	4.9 (0.1 to 9.9)	0.80 (0.70 to 0.91)	0.89 (0.74 to 1.1)
P value	.01	.01	.04	.001	.2
Adjusted results					
95% CI	11 (2.8 to 19.1)	10.1 (2.0 to 18.1)	5.3 (1.1 to 9.7)	0.81 (0.71 to 0.92)	0.87 (0.72 to 1.1)
P value	.01	.01	.01	.002	.14
Starting morning census					
95% CI	–0.1 (–3.0 to 2.8)	0.1 (–2.9 to 3.1)	1.9 (0.5 to 3.4)	0.99 (0.95 to 1.0)	0.98 (0.92 to 1.1)
P value	.96	.95	.01	.75	.63
Staffed with an APP[d]					
95% CI	–20.1 (–36.1 to –4.0)	–11.7 (–27.6 to 4.3)	1.9 (–6.2 to 10.7)	1.3 (1.0, 1.7)	1.25 (0.86 to 1.8)
P value	.01	.15	.66	.05	.25

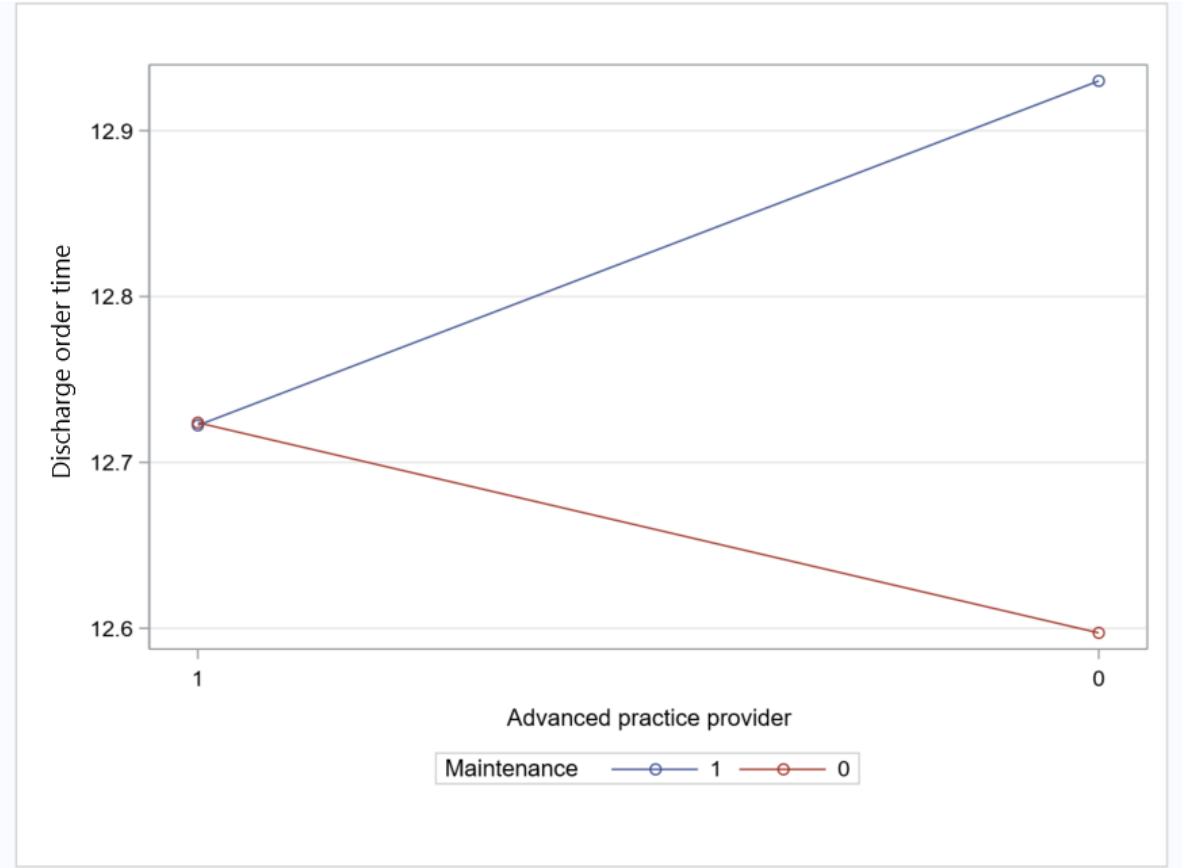
[a]Mean difference (in minutes) calculated.

[b]Mean percentage decrease calculated.

[c]Odds ratio calculated.

[d]APP: advanced practice provider.

Figure 5. Discharge order time: interaction between team staffed with an advanced practice provider and intervention period (pilot implementation vs postimplementation).



Discussion

Principal Findings

The important findings of this work are as follows:

1. There was considerable uptake and use of the Discharge Today tool for the duration of the study period, with most clinicians adding it to their patient lists in the EHR and providing discharge updates for most patients.
2. The surveyed providers and clinical staff reported that the tool was efficient to use, did not adversely affect their workflow, and was helpful for patient discharge management.
3. After adding teams staffed with an APP as an effect modifier, for teams that included an APP, there was a significant reduction in the time of day the patient left the hospital beyond the reduction seen for teams without an APP.

Other studies have described similar tools, such as the Red, Yellow, or Green Discharge tool [50] and the Kanban web-based application [51]. However, these tools were not integrated into the patient worklist, an EHR workspace that is commonly used across clinical staff, including physicians, APPs, residents, medical personnel, nurses, physical and occupational therapists, care managers and social workers, and pharmacists, thus enhancing the real time, multidisciplinary communication about discharge readiness. Recently, a similar tool was described in the pediatric setting, which was associated with an increase in the proportion of patients discharged before noon [52]. However, unlike our Discharge Today tool, this tool did not allow providers to document any tasks or clinical care required before the patient could be discharged. In addition, our Discharge Today tool allowed providers to note what time of day (before 11 AM, before 2 PM, or after 2 PM) a patient might be discharged.

Previous work has shown that hospital census and census on teams can affect the overall flow of hospitals [21,46]. In this pilot study, although the primary outcomes evaluated were nonsignificant in analysis, interactions between the number of patients assigned to a team in the morning or teams staffed with an APP and the intervention suggest that there may be effect modification at work such that the intervention is effective in certain subgroups or under certain conditions. After including starting morning census per team as an effect modifier, although nonsignificant, there was a reduction in the time of day the discharge order was entered into the EHR by the discharging physician and in the time of day the patient left the hospital. In addition, when teams were staffed with an APP, the use of this tool was associated with significantly earlier discharges beyond that seen for teams without an APP. Research has shown that discharging patients just 1 hour earlier alleviates hospital crowding and reduces access blocking [23,24]. Although we were not able to achieve the goal of discharging patients an hour earlier on average, the incremental gains from multiple solutions implemented across many patients may be additive to moving discharge times and could result in improvements in patient flow and hospital capacity.

Finally, during the maintenance period, when teams were staffed with an APP, discharge orders were entered significantly earlier by the discharging providers. Our APPs were early adopters of the tool and continue to be heavy users, which may have produced the observed improvements. We believe that these findings highlight the importance of APPs in the success of discharge initiatives. Although other studies have suggested that a multidisciplinary approach will improve the early discharge of patients [22,53,54], our study specifically investigates the effects of APP involvement. Our study suggests that APPs may be vital partners in work undertaken to improve the discharge process in an adult medicine population. A pilot study of a multidisciplinary team led by an APP and staffed by a pharmacist and nurse demonstrated a significant improvement in discharge times for patients seen by this team [55]. Similarly, previous research has shown that most providers do not prioritize discharges first as they are tending to other patients [21]; thus, using a team approach to patient care may be advantageous when working to improve throughput metrics.

Our results suggest that some effects of the tool continued even after robust stakeholder engagement efforts were reduced to periodic reminders. We observed a significant reduction in the mean time of day the physician entered the discharge order when a team was staffed with an APP over the reduction observed when a team was not staffed with an APP during the postimplementation period compared with the pilot implementation period. During this time, the hospital medicine triagist, an APP-staffed position, started using the tool for bed management, suggesting that APP use may have become more deliberate. Sustained improvement after demonstrating the effectiveness of an intervention is not often evaluated, likely because of constraints of time and available budget [56]; however, without consideration of the relevant contextual factors, evaluating whether an intervention has resulted in sustainable improvement may prove elusive [57].

Finally, our tool had high adoption and use rates, with relatively minimal incentives to do so. There were several features of our project that helped to improve adoption and use. Our stakeholder engagement process—both preimplementation and during the pilot implementation period—was robust, resulting in a product that was developed for and by frontline staff members and clinicians. In addition, the Discharge Today tool was integrated into the current workflow (ie, EHR worklists) and color coded, which serves as a visual prompt for both clinicians and frontline staff to use.

Given that this was a pilot study of this tool aiming to evaluate the user-centered design approach taken, adoption of the tool, and effectiveness in a sample of providers delivering care to hospitalized patients, the tool had not been fully scaled up across hospital settings, thus potentially limiting the effectiveness. Although our tool had high adoption rates with our target populations, it was challenging to fully implement it across all care teams across an entire hospital, and thus it took some time to scale. Since this pilot, an initiative to use the existing EHR applications to better support patient flow has been launched. As an aspect of this work, the Discharge Today tool has been integrated with other EHR functionalities to capture patients' progress toward discharge, and any roadblocks to discharge

were implemented. We suspect that as adoption continues and additional features are added, adoption will further scale, and perhaps larger effects on the desired outcomes could be seen. EHR tools intended to change clinician behavior require continuous iterative optimization and evaluation to realize their full potential.

Our study has a number of strengths. First, we describe a novel tool to communicate discharge readiness in real time to key stakeholders. Second, we had remarkable engagement by our clinicians and frontline staff members, with high use rates and overall positive feedback. Third, although our study was conducted at a single center, our sample included >4000 patients, almost 60 physicians, >40 APPs, almost 90 residents and medical students, and >160 frontline staff members during the pilot implementation period. Fourth, we have accounted in statistical modeling for the contextual factors that we hypothesized a priori could influence the effectiveness of the tool by including effect modifiers for the number of patients assigned to a team in the morning and a team staffed with an APP.

Our study also has several limitations. First, it was performed at a single university-affiliated academic hospital and was a quality improvement initiative using a pre-post study design; therefore, the results might not be generalizable to other types of institutions or other patient populations. Second, throughout our study period, we continued to optimize the tool, and thus the full effect of the tool may not have been realized at the end of the pilot implementation period. Third, for this analysis, we assessed both the discharge order time and the discharge time; however, we did not evaluate the circumstances around that gap (ie, when the patient was actually ready for discharge and any reasons for delays between the time the order was entered and the time the patient could leave). Future analyses would benefit from assessing whether the use of the Discharge Today tool closes the gap between when the discharge order is entered and when the patient is actually discharged. Fourth, there are most likely unknown confounders at work that we did not identify or include as adjustment factors. Fifth, although we did ask the providers to update the tool first thing in the morning and throughout the day as patient statuses changed, we did not ask that they otherwise change their workflow. Before, during, and after this study, there have been consistent institutional efforts asking providers to prioritize discharges first. It is possible that by asking providers to update the status, that alone could have resulted in improved discharge times regardless of the tool used; however, even with discharge-before-noon initiatives implemented at most places, <10% of hospitalists typically round on discharges first [21]. On the basis of previous literature and mixed successes around early-discharge initiatives, we believe that a multipronged approach is likely needed, including ensuring reasonable workloads, optimizing care team models, and improving communication processes [4,35,50,58]. This tool offers a potential component that is minimally intrusive and communicates across disciplines.

Finally, we were unable to account for other initiatives (eg, huddles held throughout the day to discuss patients who may be able to be discharged) intended to improve discharge times and lengths of stay that were taking place concurrently with our Discharge Today tool implementation.

Conclusions

We have described a unique, EHR-based approach to improving communication around discharge in real time with all care team members, regardless of their physical location in the hospital, that improves discharge times and lengths of stay. The Discharge Today tool allows for real time documentation and sharing of discharge statuses, and our results suggest that the tool may be useful for improving discharge times and lengths of stay, particularly if a team is staffed with an APP or possibly in higher-census situations.

Acknowledgments

The authors report funding from the Data Science to Patient Value program at the University of Colorado, Anschutz Medical Campus.

Conflicts of Interest

None declared.

References

1. Wertheimer B, Jacobs RE, Iturrate E, Bailey M, Hochman K. Discharge before noon: effect on throughput and sustainability. J Hosp Med 2015 Oct;10(10):664-669. [doi: 10.1002/jhm.2412] [Medline: 26126432]
2. White BA, Biddinger PD, Chang Y, Grabowski B, Carignan S, Brown DF. Boarding inpatients in the emergency department increases discharged patient length of stay. J Emerg Med 2013 Jan;44(1):230-235. [doi: 10.1016/j.jemermed.2012.05.007] [Medline: 22766404]
3. Beck MJ, Gosik K. Redesigning an inpatient pediatric service using Lean to improve throughput efficiency. J Hosp Med 2015 Apr;10(4):220-227. [doi: 10.1002/jhm.2300] [Medline: 25483409]
4. Wertheimer B, Jacobs RE, Bailey M, Holstein S, Chatfield S, Ohta B, et al. Discharge before noon: an achievable hospital goal. J Hosp Med 2014 Apr;9(4):210-214. [doi: 10.1002/jhm.2154] [Medline: 24446232]

5. McGowan JE, Truwit JD, Cipriano P, Howell RE, VanBree M, Garson A, et al. Operating room efficiency and hospital capacity: factors affecting operating room use during maximum hospital census. J Am Coll Surg 2007 May;204(5):865-71; discussion 871. [doi: 10.1016/j.jamcollsurg.2007.01.052] [Medline: 17481500]
6. Khanna S, Boyle J, Good N, Lind J. Early discharge and its effect on ED length of stay and access block. Stud Health Technol Inform 2012;178:92-98. [Medline: 22797025]
7. Khanna S, Boyle J, Good N, Lind J. Impact of admission and discharge peak times on hospital overcrowding. Stud Health Technol Inform 2011;168:82-88. [Medline: 21893915]
8. Khanna S, Sier D, Boyle J, Zeitz K. Discharge timeliness and its impact on hospital crowding and emergency department flow performance. Emerg Med Australas 2016 Apr;28(2):164-170. [doi: 10.1111/1742-6723.12543] [Medline: 26845068]
9. Sikka R, Mehta S, Kaucky C, Kulstad EB. ED crowding is associated with an increased time to pneumonia treatment. Am J Emerg Med 2010 Sep;28(7):809-812. [doi: 10.1016/j.ajem.2009.06.023] [Medline: 20837259]
10. Coil CJ, Flood JD, Belyeu BM, Young P, Kaji AH, Lewis RJ. The effect of emergency department boarding on order completion. Ann Emerg Med 2016 Jun;67(6):730-6.e2. [doi: 10.1016/j.annemergmed.2015.09.018] [Medline: 26655566]
11. Kulstad EB, Sikka R, Sweis RT, Kelley KM, Rzechula KH. ED overcrowding is associated with an increased frequency of medication errors. Am J Emerg Med 2010 Mar;28(3):304-309. [doi: 10.1016/j.ajem.2008.12.014] [Medline: 20223387]
12. Pines JM, Localio AR, Hollander JE, Baxt WG, Lee H, Phillips C, et al. The impact of emergency department crowding measures on time to antibiotics for patients with community-acquired pneumonia. Ann Emerg Med 2007 Nov;50(5):510-516. [doi: 10.1016/j.annemergmed.2007.07.021] [Medline: 17913298]
13. Rojas-García A, Turner S, Pizzo E, Hudson E, Thomas J, Raine R. Impact and experiences of delayed discharge: a mixed-studies systematic review. Health Expect 2018 Feb;21(1):41-56 [FREE Full text] [doi: 10.1111/hex.12619] [Medline: 28898930]
14. Gaieski DF, Agarwal AK, Mikkelsen ME, Drumheller B, Cham Sante S, Shofer FS, et al. The impact of ED crowding on early interventions and mortality in patients with severe sepsis. Am J Emerg Med 2017 Jul;35(7):953-960. [doi: 10.1016/j.ajem.2017.01.061] [Medline: 28233644]
15. Mathews KS, Durst MS, Vargas-Torres C, Olson AD, Mazumdar M, Richardson LD. Effect of emergency department and ICU occupancy on admission decisions and outcomes for critically ill patients. Crit Care Med 2018 May;46(5):720-727 [FREE Full text] [doi: 10.1097/CCM.0000000000002993] [Medline: 29384780]
16. Singer A, Thode HJ, Viccellio P, Pines J. The association between length of emergency department boarding and mortality. Acad Emerg Med 2011 Dec;18(12):1324-1329 [FREE Full text] [doi: 10.1111/j.1553-2712.2011.01236.x] [Medline: 22168198]
17. Forster A, Stiell I, Wells G, Lee A, van Walraven C. The effect of hospital occupancy on emergency department length of stay and patient disposition. Acad Emerg Med 2003 Feb;10(2):127-133 [FREE Full text] [doi: 10.1111/j.1553-2712.2003.tb00029.x] [Medline: 12574009]
18. Everall AC, Guilcher SJ, Cadel L, Asif M, Li J, Kuluski K. Patient and caregiver experience with delayed discharge from a hospital setting: a scoping review. Health Expect 2019 Oct;22(5):863-873 [FREE Full text] [doi: 10.1111/hex.12916] [Medline: 31099969]
19. Landeiro F, Roberts K, Gray A, Leal J. Delayed hospital discharges of older patients: a systematic review on prevalence and costs. Gerontologist 2019 Mar 14;59(2):e86-e97. [doi: 10.1093/geront/gnx028] [Medline: 28535285]
20. Patel H, Morduchowicz S, Mourad M. Using a systematic framework of interventions to improve early discharges. Jt Comm J Qual Patient Saf 2017 Apr;43(4):189-196. [doi: 10.1016/j.jcjq.2016.12.003] [Medline: 28325207]
21. Zoucha J, Hull M, Keniston A, Mastalerz K, Quinn R, Tsai A, et al. Barriers to early hospital discharge: a cross-sectional study at five academic hospitals. J Hosp Med 2018 Dec;13(12):816-822. [doi: 10.12788/jhm.3074] [Medline: 30496327]
22. Patel H, Yirdaw E, Yu A, Slater L, Perica K, Pierce RG, et al. Improving early discharge using a team-based structure for discharge multidisciplinary rounds. Prof Case Manag 2019;24(2):83-89. [doi: 10.1097/NCM.0000000000000318] [Medline: 30688821]
23. Powell ES, Khare RK, Venkatesh AK, Van Roo BD, Adams JG, Reinhardt G. The relationship between inpatient discharge timing and emergency department boarding. J Emerg Med 2012 Feb;42(2):186-196. [doi: 10.1016/j.jemermed.2010.06.028] [Medline: 20888163]
24. Khanna S, Boyle J, Good N, Lind J. Unravelling relationships: hospital occupancy levels, discharge timing and emergency department access block. Emerg Med Australas 2012 Oct;24(5):510-517. [doi: 10.1111/j.1742-6723.2012.01587.x] [Medline: 23039292]
25. Sharma A, Lo V, Lapointe-Shaw L, Soong C, Wu PE, Wu RC. A time-motion study of residents and medical students performing patient discharges from general internal medicine wards: a disjointed, interrupted process. Intern Emerg Med 2017 Sep;12(6):789-798. [doi: 10.1007/s11739-017-1654-4] [Medline: 28349373]
26. Smith CN, Quan SD, Morra D, Rossos PG, Khatibi H, Lo V, et al. Understanding interprofessional communication: a content analysis of email communications between doctors and nurses. Appl Clin Inform 2012 Feb 1;3(1):38-51 [FREE Full text] [doi: 10.4338/ACI-2011-11-RA-0067] [Medline: 23616899]

27. Martinez DA, Kane EM, Jalalpour M, Scheulen J, Rupani H, Toteja R, et al. An electronic dashboard to monitor patient flow at the Johns Hopkins hospital: communication of key performance indicators using the Donabedian model. J Med Syst 2018 Jun 18;42(8):133. [doi: 10.1007/s10916-018-0988-4] [Medline: 29915933]
28. Rolls D, Khanna S, Lloyd N, Reeson A, Jayasena R, McCormick C, et al. Before-after evaluation of patient length of stay in a rehabilitation context following implementation of an electronic patient journey board. Int J Med Inform 2020 Feb;134:104042 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.104042] [Medline: 31855847]
29. Walsh CM, Coldiron JS. Enhanced length of stay management through monitoring of discharge planning parameters. Am J Med Qual 1993;8(3):128-133. [doi: 10.1177/0885713X9300800304] [Medline: 8219874]
30. Clark KW, Moller S, O'Brien L. Electronic patient journey boards a vital piece of the puzzle in patient flow. Aust Health Rev 2014 Jun;38(3):259-264. [doi: 10.1071/AH13192] [Medline: 24814229]
31. Hamline MY, Rutman L, Tancredi DJ, Rosenthal JL, University of California Davis Children's Hospital Discharge Quality Improvement Working Group. An iterative quality improvement process improves pediatric ward discharge efficiency. Hosp Pediatr 2020 Mar;10(3):214-221 [FREE Full text] [doi: 10.1542/hpeds.2019-0158] [Medline: 32051222]
32. Garg T, Lee JY, Evans KH, Chen J, Shieh L. Development and evaluation of an electronic health record-based best-practice discharge checklist for hospital patients. Jt Comm J Qual Patient Saf 2015 Mar;41(3):126-131. [doi: 10.1016/s1553-7250(15)41017-7] [Medline: 25977128]
33. Tyler A, Boyer A, Martin S, Neiman J, Bakel LA, Brittan M. Development of a discharge readiness report within the electronic health record-A discharge planning tool. J Hosp Med 2014 Aug;9(8):533-539. [doi: 10.1002/jhm.2212] [Medline: 24825848]
34. Moo-Young J, Sylvester F, Dancel R, Galin S, Troxler H, Bradford K. Impact of a quality improvement initiative to optimize the discharge process of pediatric gastroenterology patients at an academic children's hospital. Pediatr Qual Saf 2019 Sep 27;4(5):e213 [FREE Full text] [doi: 10.1097/pq9.0000000000000213] [Medline: 31745516]
35. Meo N, Paul E, Wilson C, Powers J, Magbual M, Miles KM. Introducing an electronic tracking tool into daily multidisciplinary discharge rounds on a medicine service: a quality improvement project to reduce length of stay. BMJ Open Qual 2018 Jul 21;7(3):e000174 [FREE Full text] [doi: 10.1136/bmjoq-2017-000174] [Medline: 30057951]
36. Maloney CG, Wolfe D, Gesteland PH, Hales JW, Nkoy FL. A tool for improving patient discharge process and hospital communication practices: the "Patient Tracker". AMIA Annu Symp Proc 2007 Oct 11;2007:493-497 [FREE Full text] [Medline: 18693885]
37. Driscoll M, Gurka D. Using the electronic medical record to enhance physician-nurse communication regarding patients' discharge status. Nurs Adm Q 2015;39(4):E31-E37. [doi: 10.1097/NAQ.0000000000000127] [Medline: 26340250]
38. Mallipudi RM, Khan Q, Mbolu GU, Alquran L, Mehta B, Allusson V. Effectiveness of implementing a predischarge order to discharge patients before 11 a.m. J Healthc Qual 2019;41(6):e77-e82. [doi: 10.1097/JHQ.0000000000000194] [Medline: 30883462]
39. Deverka PA, Lavallee DC, Desai PJ, Esmail LC, Ramsey SD, Veenstra DL, et al. Stakeholder participation in comparative effectiveness research: defining a framework for effective engagement. J Comp Eff Res 2012 Mar;1(2):181-194 [FREE Full text] [doi: 10.2217/cer.12.7] [Medline: 22707880]
40. Chokshi SK, Mann DM. Innovating from within: a process model for user-centered digital development in academic medical centers. JMIR Hum Factors 2018 Dec 19;5(4):e11048 [FREE Full text] [doi: 10.2196/11048] [Medline: 30567688]
41. Gilmartin H, Lawrence E, Leonard C, McCreight M, Kelley L, Lippmann B, et al. Brainwriting premortem: a novel focus group method to engage stakeholders and identify preimplementation barriers. J Nurs Care Qual 2019;34(2):94-100 [FREE Full text] [doi: 10.1097/NCQ.0000000000000360] [Medline: 30148746]
42. Weir CR, Nebeker JJ, Hicken BL, Campo R, Drews F, Lebar B. A cognitive task analysis of information management strategies in a computerized provider order entry environment. J Am Med Inform Assoc 2007;14(1):65-75 [FREE Full text] [doi: 10.1197/jamia.M2231] [Medline: 17068345]
43. Keniston A, McBeth L, Pell J, Bowden K, Ball S, Stoebner K, et al. Development and implementation of a multidisciplinary electronic discharge readiness tool: user-centered design approach. JMIR Hum Factors 2021 Apr 23;8(2):e24038 [FREE Full text] [doi: 10.2196/24038] [Medline: 33890860]
44. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
45. Ford C. Interpreting log transformations in a linear model. University of Virginia Library. 2018 Aug 17. URL: https://data.library.virginia.edu/interpreting-log-transformations-in-a-linear-model/ [accessed 2020-05-04]
46. Elliott DJ, Young RS, Brice J, Aguiar R, Kolm P. Effect of hospitalist workload on the quality and efficiency of care. JAMA Intern Med 2014 May;174(5):786-793. [doi: 10.1001/jamainternmed.2014.300] [Medline: 24686924]
47. Hiza E, Gottschalk M, Umpierrez E, Bush P, Reisman W. Effect of a dedicated orthopaedic advanced practice provider in a level I trauma center: analysis of length of stay and cost. J Orthop Trauma 2015 Jul;29(7):e225-e230. [doi: 10.1097/BOT.0000000000000261] [Medline: 25463429]

48. Kleinpell RM, Grabenkort WR, Kapu AN, Constantine R, Sicoutris C. Nurse practitioners and physician assistants in acute and critical care: a concise review of the literature and data 2008-2018. Crit Care Med 2019 Oct;47(10):1442-1449 [FREE Full text] [doi: 10.1097/CCM.0000000000003925] [Medline: 31414993]
49. Kreeftenberg HG, Pouwels S, Bindels AJ, de Bie A, van der Voort PH. Impact of the advanced practice provider in adult critical care: a systematic review and meta-analysis. Crit Care Med 2019 May;47(5):722-730. [doi: 10.1097/CCM.0000000000003667] [Medline: 30720539]
50. Mathews KS, Corso P, Bacon S, Jenq GY. Using the red/yellow/green discharge tool to improve the timeliness of hospital discharges. Jt Comm J Qual Patient Saf 2014 Jun;40(6):243-252 [FREE Full text] [doi: 10.1016/s1553-7250(14)40033-3] [Medline: 25016672]
51. Rocha H, Santos A, Alcântara AC, Lima CS, Rocha SG, Cardoso R, et al. Bed management team with Kanban web-based application. Int J Qual Health Care 2018 Nov 01;30(9):708-714. [doi: 10.1093/intqhc/mzy108] [Medline: 29767742]
52. Perry M, Macias C, Chaparro J, Heacock A, Jackson K, Bode R. Improving early discharges with an electronic health record discharge optimization tool. Pediatr Qual Saf 2020 May 18;5(3):e301 [FREE Full text] [doi: 10.1097/pq9.0000000000000301] [Medline: 32607458]
53. Kane M, Weinacker A, Arthofer R, Seay-Morrison T, Elfman W, Ramirez M, et al. A multidisciplinary initiative to increase inpatient discharges before noon. J Nurs Adm 2016 Dec;46(12):630-635. [doi: 10.1097/NNA.0000000000000418] [Medline: 27851703]
54. Durvasula R, Kayihan A, Del Bene S, Granich M, Parker G, Anawalt BD, et al. A multidisciplinary care pathway significantly increases the number of early morning discharges in a large academic medical center. Qual Manag Health Care 2015;24(1):45-51. [doi: 10.1097/QMH.0000000000000049] [Medline: 25539490]
55. Bowden K, Keniston A, McBeth L, Perica K, Burden M. Discharge today: the efficacy of a multi-disciplinary discharge team. J Gen Intern Med 2021:1-4 [FREE Full text] [doi: 10.1007/s11606-021-07043-6]
56. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement Sci 2013 Oct 02;8:117 [FREE Full text] [doi: 10.1186/1748-5908-8-117] [Medline: 24088228]
57. Silver SA, McQuillan R, Harel Z, Weizman AV, Thomas A, Nesrallah G, et al. How to sustain change and support continuous quality improvement. Clin J Am Soc Nephrol 2016 May 06;11(5):916-924 [FREE Full text] [doi: 10.2215/CJN.11501015] [Medline: 27016498]
58. Statile AM, Schondelmeyer AC, Thomson JE, Brower LH, Davis B, Redel J, et al. Improving discharge efficiency in medically complex pediatric patients. Pediatrics 2016 Aug;138(2):e20153832. [doi: 10.1542/peds.2015-3832] [Medline: 27412640]

Abbreviations

APP: advanced practice provider
EHR: electronic health record
REDCap: Research Electronic Data Capture

Edited by A Kushniruk; submitted 28.01.21; peer-reviewed by R Marshall, M Esdar; comments to author 28.03.21; revised version received 23.05.21; accepted 06.06.21; published 08.11.21.

Please cite as:
Keniston A, McBeth L, Pell J, Bowden K, Metzger A, Nordhagen J, Anthony A, Rice J, Burden M
The Effectiveness of a Multidisciplinary Electronic Discharge Readiness Tool: Prospective, Single-Center, Pre-Post Study
JMIR Hum Factors 2021;8(4):e27568
URL: https://humanfactors.jmir.org/2021/4/e27568
doi:10.2196/27568
PMID:34747702

Original Paper

Computer-Aided Screening of Autism Spectrum Disorder: Eye-Tracking Study Using Data Visualization and Deep Learning

Federica Cilia[1*], PhD; Romuald Carette[2*], PhD; Mahmoud Elbattah[2*], PhD; Gilles Dequen[2*], PhD; Jean-Luc Guérin[2*], PhD; Jérôme Bosche[2*], PhD; Luc Vandromme[3*], PhD; Barbara Le Driant[1*], PhD

[1]UR-UPJV 7273, Centre de Recherche en Psychologie - Cognition, Psychisme, Organisations, Université de Picardie Jules Verne, Amiens, France

[2]UR-UPJV 4290, Modélisation, Information & Systèmes, Université de Picardie Jules Verne, Amiens, France

[3]UR-UPJV 7516, Chirurgie et Extrémité Céphalique Caractérisation Morphologique et Fonctionnelle, Université de Picardie Jules Verne, Amiens, France

[*]all authors contributed equally

Corresponding Author:
Federica Cilia, PhD
UR-UPJV 7273
Centre de Recherche en Psychologie - Cognition, Psychisme, Organisations
Université de Picardie Jules Verne
10 rue des Français Libres
Amiens, 80000
France
Phone: 33 322 827 397
Email: federica.cilia@u-picardie.fr

Abstract

Background: The early diagnosis of autism spectrum disorder (ASD) is highly desirable but remains a challenging task, which requires a set of cognitive tests and hours of clinical examinations. In addition, variations of such symptoms exist, which can make the identification of ASD even more difficult. Although diagnosis tests are largely developed by experts, they are still subject to human bias. In this respect, computer-assisted technologies can play a key role in supporting the screening process.

Objective: This paper follows on the path of using eye tracking as an integrated part of screening assessment in ASD based on the characteristic elements of the eye gaze. This study adds to the mounting efforts in using eye tracking technology to support the process of ASD screening

Methods: The proposed approach basically aims to integrate eye tracking with visualization and machine learning. A group of 59 school-aged participants took part in the study. The participants were invited to watch a set of age-appropriate photographs and videos related to social cognition. Initially, eye-tracking scanpaths were transformed into a visual representation as a set of images. Subsequently, a convolutional neural network was trained to perform the image classification task.

Results: The experimental results demonstrated that the visual representation could simplify the diagnostic task and also attained high accuracy. Specifically, the convolutional neural network model could achieve a promising classification accuracy. This largely suggests that visualizations could successfully encode the information of gaze motion and its underlying dynamics. Further, we explored possible correlations between the autism severity and the dynamics of eye movement based on the maximal information coefficient. The findings primarily show that the combination of eye tracking, visualization, and machine learning have strong potential in developing an objective tool to assist in the screening of ASD.

Conclusions: Broadly speaking, the approach we propose could be transferable to screening for other disorders, particularly neurodevelopmental disorders.

(JMIR Hum Factors 2021;8(4):e27706) doi:10.2196/27706

KEYWORDS

autism spectrum disorder; screening; eye tracking; data visualization; machine learning; deep learning; AI; ASS; artificial intelligence; ML; screening; adolescent; diagnosis

Introduction

ASD Characteristics

Autism spectrum disorder (ASD) has been described as a pervasive developmental disorder characterized by a set of impairments including social communication problems, difficulties with reciprocal social interactions, and unusual patterns of repetitive behaviors or interests [1]. During naturalistic interaction, making and maintaining eye contact is not always easy or spontaneous for ASD-diagnosed individuals. Such troubling deficits can unfortunately place a considerable strain on their lives and their families. Nevertheless, these disturbances are not better explained by intellectual disability or global developmental delay [1].

Early diagnosis may lead to early intervention, which generally proves beneficial for both the child and the family. The diagnosis process usually involves a set of tests that can require hours of clinical examinations or is based on an interview with the parents. Furthermore, the variation of symptoms with regard to deficits in social communication and social interaction as well as the social communication impairments and restricted, repetitive patterns of behavior make the identification of ASD more complicated to decide. In this respect, computer-aided technologies have been embraced to provide helpful guidance through the course of examination and assessment. Examples include magnetic resonance imaging, electroencephalography [2], and eye tracking, which will be considered in this study. Eye-tracking technology has received particular attention in the ASD context since abnormalities of eye gaze have been consistently recognized as the hallmark of autism in general [3,4]. A considerable number of other psychology studies in eye tracking have been based on the particularities of eye movements in response to verbal or visual cues as signs of ASD [5-7]. In particular, these studies have highlighted social-related difficulties in children with ASD, especially when face stimuli are used (eg, in a face-to-butterfly categorical visual search task [8] and unsuitable extraction of visual information via eye fixations for emotion recognition [9]).

This study provides a meeting point for eye tracking and machine learning (ML) for supporting the diagnosis of ASD. It is part of an interdisciplinary collaboration between research units of psychology and artificial intelligence at the University of Picardy Jules Verne in France. Our approach is distinctively based on the premise that visual representations of eye-tracking recordings can effectively serve as features for discriminating individuals diagnosed with ASD. At its core, the key idea is to compactly render eye movements into an image-based format while maintaining the dynamic characteristics of eye motion (eg, velocity) using color gradients. In this manner, diagnostic-related tasks can be approached as a problem of image classification or analysis. The applicability of the proposed approach will be demonstrated based on the classification accuracy. Further, we will support our results with a statistical analysis that will explore possible correlations between the Childhood Autism Rating Scale (CARS) [10] and the dynamics of eye movement among participants diagnosed with ASD.

Eye Tracking for ASD Screening

Eye tracking has been used in numerous research studies. It can be described as the process of capturing, tracking, and measuring eye movements or the absolute point of gaze (POG), which refers to the point where the eye gaze is focused in the visual scene [11]. The significance of such technology is that it allows for an objective and quantitative method of recording the viewer's POG. The interpretation of eye movement can be effectively used in interactive applications or for diagnostic purposes.

Eye trackers aim to capture 3 basic categories of eye movements: (1) fixation, (2) saccade, and (3) blink. A fixation is the brief moment that occurs while pausing the gaze on an object so that the brain can perform the perception process. The average duration of fixation typically ranges from 150 ms to 300 ms [12]. However, the fixation duration is dependent on the context. The duration of our fixations differs when we are reading on paper (230 ms) or on a screen (553 ms) [13], or when we are watching a naturalistic scene on a computer (330 ms) [14]. Further, accurate perception requires constant scanning of the object with rapid eye movements, which are called saccades. Saccades include quick, ballistic jumps that take about 30-120 ms each [15]. On the other hand, a blink is often a sign that the system has lost track of the eye gaze. Eye-tracking scanpaths have been commonly used as a practical means for depicting gaze behavior in a visual manner. A scanpath represents a sequence of consecutive fixations and saccades as a trace through time and space and may overlap with itself [16].

Abundant studies have sought to take advantage of eye-tracking applications for studying and analyzing eyes movements. For instance, a team of psychologists and neuroscientists recently showed that children with ASD have faster eye movements than do children with typical development, but these results depend on the visual task the children are asked to perform. If they are faster while remaining precise in prosaccade tasks with a gap paradigm, the same children are less accurate but faster than are children with typical development in another gap paradigm during short visual search. This means that children with ASD favor speed over accuracy and that they have shorter saccadic latencies [8]. Moreover, Vabalas and Freeth [17] demonstrated that in face-to-face interactions, eye movements were different among individuals depending on where they fell on the autism spectrum. Specifically, persons with high autistic traits were observed to experience shorter and less frequent saccades. Conversely, Liberati et al [18] showed greater saccade amplitude and higher frequency in children with ASD than did control children. However, in Liberati et al's study, the Tobii Eye Tracker used had a sampling rate of 60 Hz, and they extracted the raw data to create clusters using k-means clustering, whereas Vabalas and Freeth used SMI eye-tracking glasses with a rate of 24 Hz. Thus, the question does not appear to be settled and seems to depend largely on the equipment and data used (eg, autistic traits according to a questionnaire vs autistic persons). In another study, eye-tracking was used to identify children diagnosed with ASD based on the duration of fixations and the number of saccades [19]. The results showed that participants with ASD spent significantly more time fixating on dynamic geometric images compared to other diagnostic groups.

Likewise, a longitudinal study examined the fixation patterns of infants from 2 to 6 months of age [20]. It was found that infants diagnosed with ASD exhibited a mean decline in fixations, which was not observed for those who did not develop ASD afterwards. Moreover, another cohort study suggested the strong potential of eye tracking as an objective tool for quantifying the risk of autism and estimating the severity of its symptoms [21]. A high diagnostic accuracy was demonstrated in this regard as well.

ML for ASD Screening

ML is subfield of computer science involved in providing computers the ability to learn without being explicitly programmed [22]. In contrast to traditional programming, ML attempts to extrapolate algorithms from data exclusively. Thus, the power of ML is that it allows for extracting insights, making predictions, or taking actions with minimal human intervention (if any). The development of ML can be broadly organized into supervised or unsupervised models. On the one hand, supervised ML deals with labeled examples, where the desired output is known precisely. The learning algorithm receives a set of inputs along with corresponding labels, and the algorithm can learn by comparing predicted labels to the actual ones. The model can be iteratively optimized to minimize error. On the other hand, unsupervised ML uses training data that do not include any output information (ie, labels). Unsupervised models (eg, clustering and association rules) can provide descriptive knowledge to help understand the inherent structure or properties of the data.

The coupling of eye tracking with ML is currently leveraging further capabilities for advancing ASD diagnosis and its applications. The literature includes several contributions in this context. For instance, Pusiol et al [23] worked on the analysis of the eye focus on the face during conversations. Their analysis was specifically applied to children with developmental disorders or those with fragile X syndrome. They tested a set of classification models, including recurrent neural networks, support vector machine, Naive Bayes, and the hidden Markov model. With recurrent neural networks, they were able to reach a high prediction accuracy of 86% and 91% for the classification of female and male fragile X syndrome, respectively. Another recent study applied ML on eye-tracking output to predict ASD [24]. The ML model included features related to the saccade eye movement (eg, amplitude, duration, and acceleration). The experiments were aimed at detecting ASD among a set of 17 children aged 8 to 10 years. Despite the use of a limited data set and a relatively simple model, the findings demonstrated the promising potential of ML for this application.

Other recent studies have focused on predicting the visual attention of children with ASD. For instance, Wei et al [25] proposed a saliency prediction model based on a convolutional neural network (CNN), but they concluded that it is necessary to first train the model on an eye-tracking data set of typical development to enable more effective saliency prediction. Jiang et al [26] proposed a method with 86% accuracy that classifies eye fixations based on a comprehensive set of features and that integrates task performance, gaze information, and facial features extracted using a deep neural network. Their work focused on a population of children with ASD between the ages of 8 and 17 years whose intellectual level was highly disparate (IQ score range 58-137).

Compared to the literature, the main distinction of this paper is that it is purely reliant on the visual representation of eye-tracking scanpaths. The study aims to produce scanpath visualizations that can represent the spatial patterns of gaze behavior and its dynamics. In this way, the vision-based approach allows for approaching the diagnosis problem as a typical task of image classification and is a continuation of our earlier work [24,27]. Our initial work applied a different set of features based on the events of fixations and saccades. We have transformed the eye-tracking data into a visual representation [27,28]. This study builds on our earlier efforts in an attempt to develop more sophisticated ML models using deep learning.

Methods

Recruitment

A group of 59 children took part in this study. It was highly desirable to have participants at an early stage of development, as the principal goal was supporting the early detection and diagnosis of ASD. Specifically, all participants were school-aged children of a mean age of about 8 years. This somewhat advanced age was indispensable here because in our region there were not enough diagnosed children younger than 6 years, and the time it takes to consult a doctor to make a diagnosis can be as long as 2 years. For the group of typically developing children (non-ASD), parental reports of any possible concerns were carefully considered.

The ASD diagnosis was confirmed by health professionals using standardized tools (Autism Diagnostic Interview-Revised [ADI-R], and Autism Diagnostic Observation Schedule–Generic [ADOS-G]). However, we did not get permission to read the children's medical files. ADI-R and ADOS-G scores were not analyzed in this study. The participants were broadly organized into 2 groups: (1) diagnosed with ASD or (2) non-ASD. Children diagnosed with ASD were examined in multidisciplinary ASD specialty clinics. The intensity of autism was estimated by psychologists using the French version of the CARS [29], while communication level was assessed with the French version of the Early Social Communication Scale (Echelle d'évaluation de la Communication Sociale Précoce [ECSP]) [30]. Table 1 summarizes the statistics of the participants.

All the children's parents or legal guardians were informed of the objectives of the study, the nature of the tasks that would be administered, and the fact that they could withdraw their agreement at any time. Their informed consent was received in writing in accordance with the Declaration of Helsinki of June 1964 (amended at the 64th General Assembly of the World Health Organization in October 2013). Moreover, all children gave their agreement to participate, and if they wished, parents could be present with their children in the experimental room. This study did not require authorization from an ethics committee based on the recommendations for psychological research in France and in agreement with the national and institutional guidelines.

Table 1. Summary of participant statistics.

Child group	ASD[a] (n=29)	Non-ASD (n=30)
Males, n	19	19
Chronological age (years, months), mean (SD)	7, 7 (2, 6)	8 (2, 8)
Developmental age on the ESCS[b] (months, days), mean (SD)	24, 10 (6, 8)	23, 15 (6, 7)
Total ECSP score, mean (SD)	141, 1 (50, 3)	139, 18 (49, 5)
CARS[c] score (minimum score=15; autism cutoff > 30), mean (SD)	32, 9 (6, 4)	15 (0)

[a]ASD: autism spectrum disorder.

[b]ECSP: Echelle d'évaluation de la Communication Sociale Précoce.

[c]CARS: Childhood Autism Rating Scale.

Apparatus and Stimuli

The SMI RED250 remote eye tracker (250 Hz, SensoMotoric Instruments) was the main instrument used to perform the eye-tracking function. The device belongs to the category of screen-based eye trackers. It can be conveniently placed at the bottom of the screen of a desktop PC or laptop. In our case, a 17-inch monitor with a 1280 x 1024 resolution was used.

Further stimuli were presented from the SMI Experiment Center software. Stimuli represented multiple distinct types used in the eye gaze literature. Examples included static and dynamic naturalistic scenes with and without receptive language, joint attention stimuli, static face or objects, and cartoon stimuli. The average duration of eye-tracking experiments was about 10 minutes. Participants were mainly examined for the quality of eye contact with the presenter and the level of focus on other elements. A 5-point calibration scheme was used. The calibration routine was followed by a set of verification procedures.

Procedure

The participants were invited to watch a set of photographs and videos, which included scenarios tailored specifically to stimulate the eye movement across the screen area. Participants could be seated on their own or on their parents' lap at an approximately 60-cm distance from the display screen. The experiments were conducted in a quiet room at the university premises. Physical white barriers were also used to reduce visual distraction.

The scenarios varied in content and length in order to allow for analysis of the ocular activity of participants from different perspectives. In general, videos were designed to include visual elements that are especially attractive to children (eg, colorful balloons and cartoons). Specifically, the stimuli presented are part of various psychological studies. One of these studies involves the presentation of 3 videos including a situation of joint attention initiation (duration of 58 seconds per video) and 18 photos from the same situation (5 seconds per photo). The scene presented in the video and corresponding photos started with an attention grabber (ie, a hand-waving cartoon). The woman in the video then said, "Hello, how are you?" to the child and looked, verbalized, pointed, and/or verbalized at a joint attention target present or absent to the children's visual field. All conversations were performed in French as the native language of participants.

The assessment of gaze following included 12 videos (4 seconds each in duration) of an actor with a neutral face first engaging in direct gaze and then shifting to 1 out of 3 objects. In 6 videos, the actor shifted his eyes and head to the target, and in 6 other videos, he only moved his eyes to the target. The same actors were engaged in another research protocol where their photo was shown for 5 seconds on half the screen next to an object. Other stimuli presented scenes with emotional valence extracted from cartoons in which the faces of the characters expressed an emotion that was either contingent or not contingent on the previous scene (total duration 5 minutes). Moreover, in all tests, the interstimulus interval lasted 2 seconds, during which a central crosshair was presented. The differences between the stimuli used included dynamic or static, human (male and female) or cartoon, and human or object. The counterbalancing of stimuli for participants and the number of participants included allowed the artificial intelligence and psychology teams to collaborate on the basis of this predefined research protocol. The results of these tests have been partly exploited, presented, and published [31-33].

Data Transformation: Visualization of Eye-Tracking Scanpaths

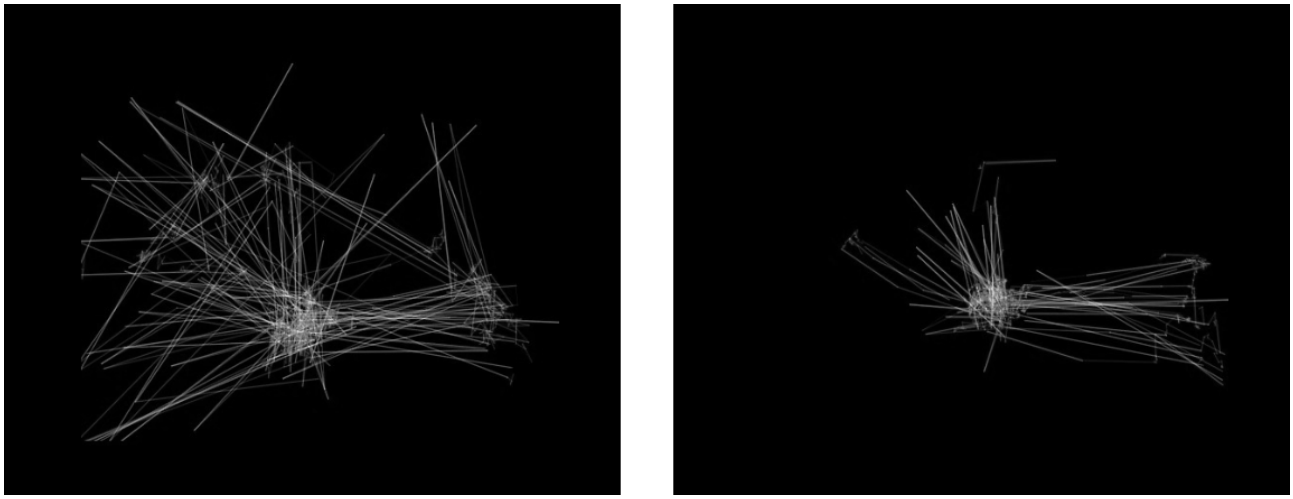
The premise of this study is based on the learning of visual patterns included in eye-tracking scanpaths. Specifically, scanpaths are used as a means to compactly describe the gaze movements into a visual representation that can simplify the learning process. Further, the scanpaths were also used to visually encode the dynamics of eye motion using color gradients. To achieve this, we used the coordinates in eye-tracking records, which represented the participant's POG during the experiment runtime. Based on the change in POG over time, we were able to calculate the velocity of gaze movement. Subsequently, the scanpath and computed dynamics were transformed into images. For each participant, a set of images was constructed in 3 steps: (1) A line was drawn for each transition from (x_t, y_t) to (x_{t+1}, y_{t+1}), where t represents a point of time during the experiment. (2) The change in color across lines was used to visualize the movement dynamics. Through use of a grayscale spectrum, the color values were tuned based on the magnitude of velocity (ie, speed) with respect

to time. (3) The images constructed were vertically mirrored since the origin was located at the bottom of the screen.

Images were constrained to contain approximately the same level of information. Specifically, a threshold was applied to limit the number of points to be drawn. The threshold was aimed to be high enough to sufficiently describe the pattern of gaze behavior. However, too-high values could increase the possibility of producing cluttered visualizations. Therefore, several tests were conducted to choose an appropriate value for the threshold. With a limit ranging from 100 to 150, images seemed to include fewer lines, which turned out to poorly discriminate the 2 classes of participants. Eventually, we decided to set the threshold to 200, which could largely strike an adequate balance and captured the key features of motion. The limit is was not a velocity threshold but a limitation to the amount of consecutive points drawn on any given scanpath image. We limited the dynamic values to a bound equal to a quarter of the diagonal of the screen because any higher movement would not be normal given the scenarios used for the capture.

The visualizations were produced using Python (Python Software Foundation) and a popular Matplotlib library [34]. The visualizations resulted in an image data set from the 59 participants who had viable data on an average of 15.19 different stimuli, allowing us to generate a total of 547 images (328 for non-ASD participants and 219 for those diagnosed with ASD), which corresponded to an average of 9.27 images per child (10.93 for non-ASD participants and 7.55 for those diagnosed with ASD). The default image dimensions were set as 640 x 480. The scanpath images were directly drawn from the raw data produced by the eye-tracking device. A more comprehensive presentation of the data set construction was elaborated upon in an earlier publication [27]. The data set was made freely available to be used by other studies investigating the potentials of eye tracking within the ASD context. Figure 1 presents 2 visualizations corresponding to participants with and without ASD.

Figure 1. Visualization of eye-tracking scanpaths. The image on the left is from a participant diagnosed with autism spectrum disorder, while the one on the right is from a participant without the disorder.



Data Preprocessing and Augmentation

Eye trackers can provide the POG coordinates on the screen. The coordinates were genuinely significant to implement our approach in terms of visualizing the gaze scanpath and computing its dynamics (eg, velocity). The eye-tracking records describe the category of movement and the POG for both eyes over time. To simplify the learning process, a set of image processing techniques was applied as follows. First, the black background was cropped from images as much as possible. The cropping was implemented using the OpenCV library. Second, all images were consistently scaled down to 256 x 256 dimensions. Resizing the images helped to reduce the problem of dimensionality by decreasing the number of features under consideration. The impact of resizing was also examined in the initial ML experiments.

Further, we applied image augmentation to produce variations of the scanpath images. Augmentation was recognized to generally improve the prediction accuracy in image classification applications [35,36]. The data set was augmented with an additional 2735 samples, where 5 synthetic samples were generated for each image. The data augmentation process was implemented using the Keras library [37], which includes an easy-to-use application programming interface for that purpose.

Classification Model

The ML work described here falls into the category of supervised learning. The basic goal was to develop a binary classifier that could predict the class of participant (ie, ASD or non-ASD) based on the scanpath images. The classification model was implemented using an artificial neural network approach. Specifically, we designed a deep CNN.

CNNs typically include 3 categories of layers including convolutional layers, pooling layers, and fully connected layers [38]. The learning process goes through a series of convolutions and pooling, which break down the input image into a set of features maps. Convolutional layers initially attempt to extract features from the image through applying a convolutional kernel all over the image. Subsequently, pooling layers work on reducing the dimensions of feature maps extracted. Eventually, the output of this process usually feeds into a fully connected layer structure to produce the final prediction. In our case, the

CNN model was composed of 4 convolutional layers, 4 pooling layers, and 2 fully connected layers. In addition, dropout layers were used, which help reduce the possibility of overfitting [39].

Results

Classification Accuracy

The classification accuracy was analyzed based on the receiver operating characteristic (ROC) curve. The ROC curve plotted the relationship between the true-positive rate and the false-positive rate across a full range of possible thresholds. Figure 2 plots the ROC curve of the CNN model. The figure also shows the approximate value of the area under the curve along with its standard deviation based on the 3-fold cross-validation. As it appears, the model could provide a notable prediction accuracy (≈90%), recall (ie, sensitivity; ≈83%), and precision (≈80%).

The model was implemented using the Keras library [37] with Python. The model was trained based on 3 rounds of cross-validation over 3 epochs. Training the model took ≈3 minutes using a single Tesla K80 GPU. Figure 3 demonstrates the model loss in training and validation over 3 epochs with 20% of the data set used for validation.

Figure 2. Receiver operating characteristic curve of the convolutional neural network model. AUC: area under the curve.

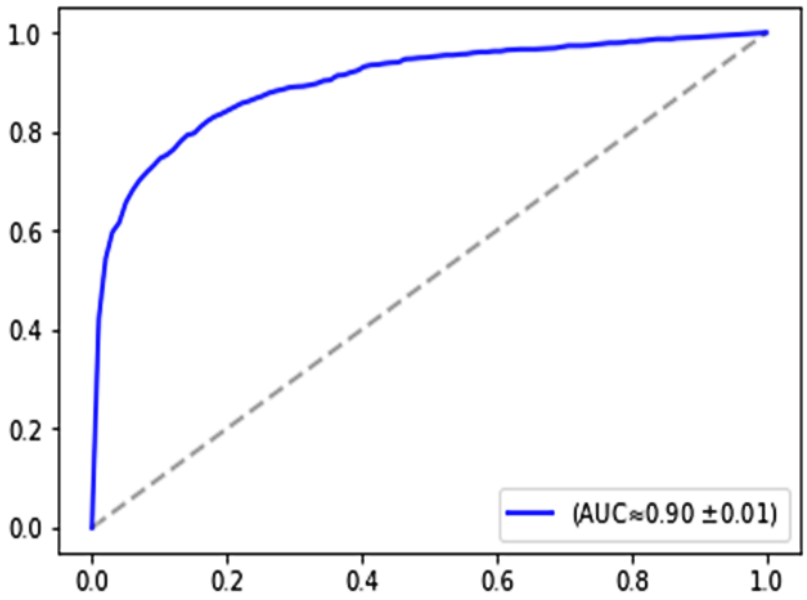
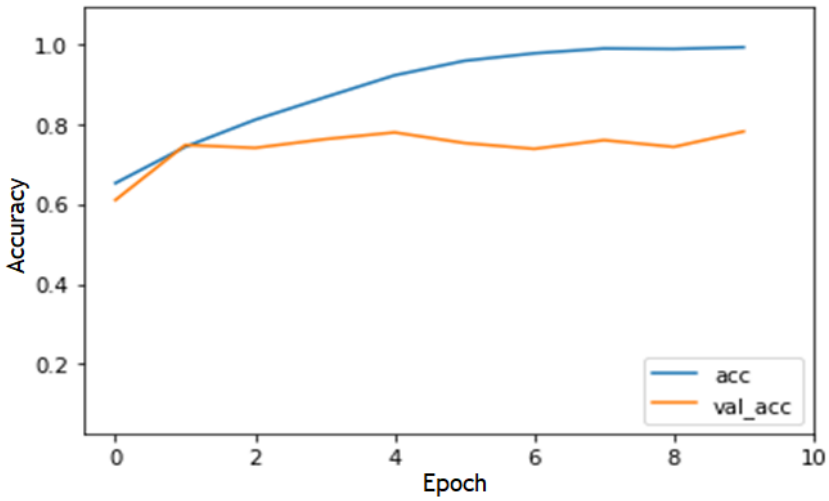


Figure 3. Model loss in the training and validation sets. acc: accuracy; val-acc: validation accuracy.

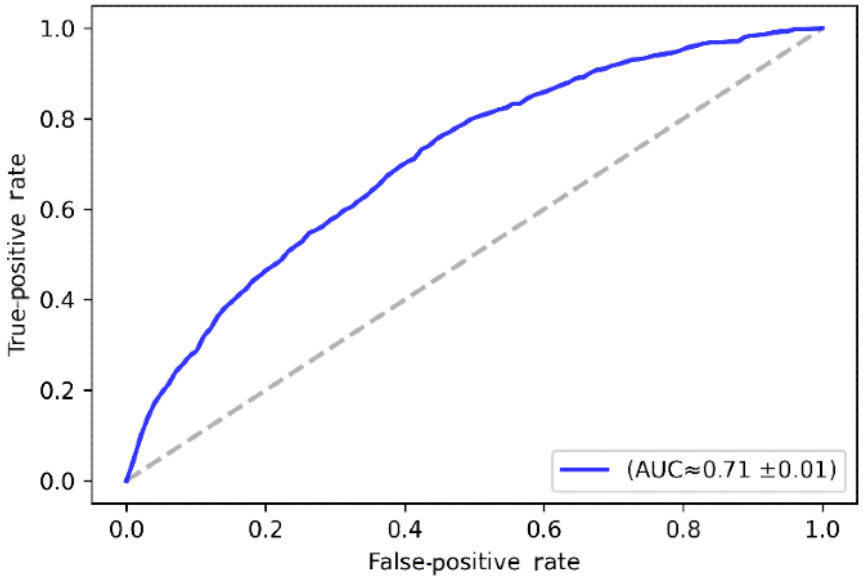


To further examine the model performance, the training and test sets were split based on the participants. The data set was split into training and test sets based on a 3-fold cross-validation using 3 stepwise procedures. First, the group of 59 participants were randomly split into 2 independent sets (ie, training and test), then the images were matched and loaded into the training and test sets based on the IDs of participants, and finally, these 2 steps were repeated for each round of the cross-validation process.

The features were extracted using the convolutional layers in the CNN model. The learning process goes through a series of convolutions and pooling, which break down the input image into a set of features maps. Convolutional layers initially attempt

to extract features from the image through applying a convolutional kernel all over the image. Subsequently, pooling layers work on reducing the dimensions of feature maps extracted. Eventually, the output of this process usually feeds into a fully connected layer structure to produce the final prediction. Expectedly, the model performance declined as shown in Figure 4. The accuracy (≈71%) could still be viewed as promising given the relatively small data set.

Figure 4. Receiver operating characteristic curve of all the data divided according to participants. AUC: area under the curve.

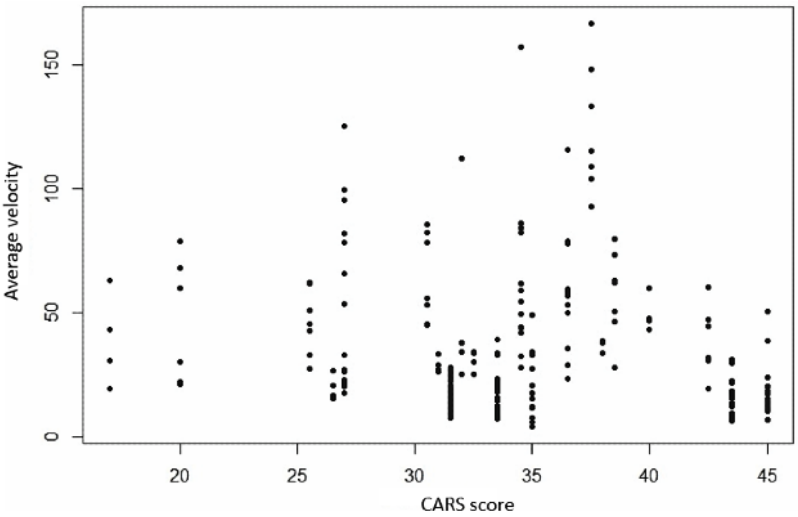


Correlation Analysis

This section serves as an integral part that supports the experimental results gained by our approach. Through statistical analysis, we attempted to explore possible correlations between the CARS score and the dynamics of eye movement in the eye-tracking scanpaths. Initially, the average velocity magnitude was calculated per image. In this way, the CARS scores of participants could be considered multiple times with respect to velocity. This could help mitigate the effect of outliers in eye-tracking experiments.

The patterns largely revealed the nonlinearity of the relationship between CARS scores and velocity. Therefore, standard correlation tests (eg, Pearson's *r*) would not be useful in such a case. Instead, we made use of the maximal information coefficient (MIC) [40]. The MIC score can describe the correlation between variable pairs regardless of a linear or nonlinear relationship. The score provided by MIC can be roughly considered as the coefficient of determination (R^2). The MIC method has been embraced in a large number of studies to find correlations in complex data sets related to, for example, biology and genomics [41,42]. We used the Minerva R package [43], which greatly facilitated the computation of MIC. The MIC values (presented in Figure 5) suggested strong correlation between CARS and velocity (MIC= 0.79). The high correlation score result could partly validate the accuracy demonstrated by the classification model, whereas the velocity was visually encoded within the scanpath images.

Figure 5. Average velocity depending on CARS value. Avg: average; CARS: Childhood Autism Rating Scale.



Demo Application

A demo application was developed to serve as a practical illustration of our approach. The application links the 3 components of eye-tracking, visualization, and ML together to support the diagnosis process of ASD.

The application goes through 3 steps as follows. First, the user is asked to upload the eye-tracking data. The data records should describe the coordinates of the viewer's gaze into the screen along with the associated time. Second, the application produces a visualization of the eye-tracking scanpath. Eventually, the application calls the prediction web service, which returns the prediction from the trained classification model. Azure ML is employed to host the classification model and the Python implementation used to produce visualizations. The application can be accessed online by asking the authors for the URL link.

Discussion

Principal Results

This study demonstrated the strong potential of eye tracking as an objective tool for assisting ASD diagnosis. Indeed, abnormal eye gaze has been a hallmark characteristic of ASD [6,7]. Over several years, eye-tracking technology has been widely used to study attention impairment among individuals diagnosed with ASD [8,9,19,20]. In this paper, we introduced an additional dimension to the representation of eye-tracking scanpaths, and we demonstrated its effectiveness for training a classification model. In similar fashion to Frazier et al [21], we used static and dynamic stimuli including social and nonsocial images. However, adding nonsocial targets may be particularly important for increasing the relationship between nonsocial attention and ASD symptoms.

The empirical results provided a set of implications to be considered. First, the ML experiments confirmed the core idea behind our approach, which hinges on the visual representation of scanpaths. The classification accuracy indicated that scanpath visualizations were able to successfully pack the information of gaze motion and its underlying dynamics. This evidently translates into the validity of employing such visual patterns in order to diagnose individuals with ASD.

Equally important, the study brought further interesting insights into the features of autistic gaze. We provided a statistical analysis that revealed possible a correlation between the level of autism (ie, CARS) and the dynamic characteristics of eye motion (eg, velocity). The analysis can lend support to the findings of Vabalas and Freeth [17], which suggested that individuals with high autistic traits tend to have shorter and less frequent saccades compared to others with low autistic traits. However, the lack of a benchmark data set in the ASD literature makes it difficult to strictly compare our results to other ML approaches. A future larger project (with a cohort of children with ASD and typical children at different ages) should be considered and should analyze the socio-cognitive and cognitive profiles of children with ASD using eye tracking. The extensive literature on these different processes may be considered in connection with the study of gaze distinctiveness in children with ASD.

Limitations

Even though the results presented in this study are promising, the following set of limitations should be highlighted. The primary limitation was the relatively small number of participants. In a future study, a data augmentation method for an ASD data set may be considered [44]. The interpretation of our results is limited by the fact that we did not have access to all the standardized test scores (ie, ADI and ASOS) used to clinically diagnose our study population. Also, the inclusion of ADI and ADOS scores could have provided further interpretation of the results. Another relevant issue of concern is the duration of video scenarios, which were relatively short. Perhaps longer scenarios might have allowed for a richer representation of the gaze behavior. Indeed, if the algorithm currently used and the age group of children in the model are limited for the moment, future work on a larger cohort will allow us to improve the study. In fact, despite limitations, we still believe this study can serve as the kernel for further interesting applications of the proposed approach.

Conclusions

To conclude, the combination of eye tracking, visualization, and ML may hold considerable potential for the development of an objective tool to assist the diagnosis of ASD. These results can be used and new data analyzed to create a screening tool for health professionals Further, features related to the dynamics of eye movement can also be considered as candidate features for developing predictive models, and recently published deep neural network methodologies can be adapted to our model [45]. Eye-tracking measures which require limited technical expertise can be quickly managed during diagnostic interviews. Moreover, parents seem to have high acceptance of eye tracking as part of the clinical evaluation because the visual results are easier to understand than is the ADI cutoff, for example. In fact, for some parents, a lack of an objective measure can lead to delayed or diminished acceptance of the clinical diagnosis. However, some limitations may still delay the clinical adoption of eye tracking as an objective measure (eg, hardware and software costs), yet these issues can be reduced by consolidating the synergy between clinical structures and academic research.

Acknowledgments

We would like to thank the participants in this study, their families, and the institutions that allowed us to work with them, as well as the health care teams for the invaluable help they gave us while we were conducting the research. We would also like to thank the HEALTHINF conference and its participating experts, whose comments in the 2019 conference have enabled us to improve our work.

This article was supported by the research quality bonus of the University of Picardie Jules Verne at Amiens.

Authors' Contributions

All authors listed have made substantial, direct, and intellectual contributions to the work, and approved it for publication.

Conflicts of Interest

None declared.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders DSM. Arlington: VA: American Psychiatric Publishing; 2013:5.
2. Bosl W, Tierney A, Tager-Flusberg H, Nelson C. EEG complexity as a biomarker for autism spectrum disorder risk. BMC Med 2011 Feb 22;9:18 [FREE Full text] [doi: 10.1186/1741-7015-9-18] [Medline: 21342500]
3. Frazier TW, Strauss M, Klingemier EW, Zetzer EE, Hardan AY, Eng C, et al. A meta-analysis of gaze differences to social and nonsocial information between individuals with and without autism Meta-Analysis of Gaze Differences to Social and Nonsocial Information Between Individuals With and Without Autism. J Am Acad Child Adolesc Psychiatry 2017 Jul;56(7):546-555 [FREE Full text] [doi: 10.1016/j.jaac.2017.05.005] [Medline: 28647006]
4. Shic F. Eye tracking as a behavioral biomarker for psychiatric conditions: the road ahead. J Am Acad Child Adolesc Psychiatry 2016 Apr;55(4):267-268. [doi: 10.1016/j.jaac.2016.02.002] [Medline: 27015716]
5. Jones E, Gliga T, Bedford R, Charman T, Johnson M. Developmental pathways to autism: a review of prospective studies of infants at risk. Neurosci Biobehav Rev 2014 Feb;39:1-33 [FREE Full text] [doi: 10.1016/j.neubiorev.2013.12.001] [Medline: 24361967]
6. Sepeta L, Tsuchiya N, Davies MS, Sigman M, Bookheimer SY, Dapretto M. Abnormal social reward processing in autism as indexed by pupillary responses to happy faces. J Neurodev Disord 2012 Jun 07;4(1):17 [FREE Full text] [doi: 10.1186/1866-1955-4-17] [Medline: 22958650]
7. Kylliäinen A, Wallace S, Coutanche MN, Leppänen JM, Cusack J, Bailey AJ, et al. Affective-motivational brain responses to direct gaze in children with autism spectrum disorder. J Child Psychol Psychiatry 2012 Jul;53(7):790-797. [doi: 10.1111/j.1469-7610.2011.02522.x] [Medline: 22276654]
8. Kovarski K, Siwiaszczyk M, Malvy J, Batty M, Latinus M. Faster eye movements in children with autism spectrum disorder. Autism Res 2019 Feb;12(2):212-224. [doi: 10.1002/aur.2054] [Medline: 30585440]
9. Król ME, Król M. A novel machine learning analysis of eye-tracking data reveals suboptimal visual information extraction from facial stimuli in individuals with autism. Neuropsychologia 2019 Jun;129:397-406. [doi: 10.1016/j.neuropsychologia.2019.04.022] [Medline: 31071324]
10. Schopler E, Van BM, Wellman G, Love S. Childhood Autism Rating Scale, Second Edition (CARS-2). Los Angeles, CA: Western Psychological Service; 2010.
11. Majaranta P, Bulling A. Eye tracking and eye-based human–computer interaction. In: Fairclough SH, Gilleade K, editors. Advances in Physiological Computing. London: Springer Publishing; 2014:39-65.
12. Tullis T, Albert B. Behavioral and physiological metrics. In: The User Experience: Collecting, Analyzing, And Presenting Usability Metrics. Amsterdam: Elsevier/Morgan Kaufmann; 2013:163-186.
13. Benedetto S, Carbone A, Pedrotti M, Le Fevre K, Bey L, Baccino T. Rapid serial visual presentation in reading: The case of Spritz. Computers in Human Behavior 2015 Apr;45:352-358. [doi: 10.1016/j.chb.2014.12.043]
14. Henderson J. Human gaze control during real-world scene perception. Trends Cogn Sci 2003 Nov;7(11):498-504. [doi: 10.1016/j.tics.2003.09.006] [Medline: 14585447]
15. Jacob R. Eye tracking in advanced interface design. In: Barfield W, Furness TA, editors. Virtual Environments and Advanced Interface Design. New York: Oxford University Press; 1995:258-288.
16. Goldberg WA, Jarvis KL, Osann K, Laulhere TM, Straub C, Thomas E, et al. Brief report: early social communication behaviors in the younger siblings of children with autism. J Autism Dev Disord 2005 Oct;35(5):657-664. [doi: 10.1007/s10803-005-0009-6] [Medline: 16167088]
17. Vabalas A, Freeth M. Brief report: patterns of eye movements in face to face conversation are associated with autistic traits: evidence from a student sample. J Autism Dev Disord 2016 Jan;46(1):305-314. [doi: 10.1007/s10803-015-2546-y] [Medline: 26249261]
18. Liberati A, Fadda R, Doneddu G, Congiu S, Javarone MA, Striano T, et al. A statistical physics perspective to understand social visual attention in autism spectrum disorder. Perception 2017 Aug;46(8):889-913. [doi: 10.1177/0301006616685976] [Medline: 28056653]
19. Pierce K, Conant D, Hazin R, Stoner R, Desmond J. Preference for geometric patterns early in life as a risk factor for autism. Arch Gen Psychiatry 2011 Jan;68(1):101-109 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2010.113] [Medline: 20819977]
20. Jones W, Klin A. Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. Nature 2013 Dec 19;504(7480):427-431 [FREE Full text] [doi: 10.1038/nature12715] [Medline: 24196715]

21. Frazier TW, Klingemier EW, Beukemann M, Speer L, Markowitz L, Parikh S, et al. Development of an objective autism risk index using remote eye tracking. J Am Acad Child Adolesc Psychiatry 2016 Apr;55(4):301-309 [FREE Full text] [doi: 10.1016/j.jaac.2016.01.011] [Medline: 27015721]
22. Samuel AL. Some studies in machine learning using the game of checkers. IBM J. Res. & Dev 2000 Jan;44(1.2):206-226. [doi: 10.1147/rd.441.0206]
23. Pusiol G, Esteva A, Hall S, Frank M, Milstein A, Fei-fei L. Vision-based classification of developmental disorders using eye-movements. 2016 Presented at: International Conference on Medical Image Computing and Computer-Assisted Intervention; Oct 17-21, 2016; Athens. [doi: 10.1007/978-3-319-46723-8_37]
24. Carette R, Cilia F, Dequen G, Bosche J, Guerin J, Vandromme L. Automatic autism spectrum disorder detection thanks to eye-tracking and neural network-based approach. Internet Things Technol Healthc Autom Springer 2018:1-7. [doi: 10.1007/978-3-319-76213-5_11]
25. Wei W, Liu Z, Huang L, Nebout A, Le MO. Saliency prediction via multi-level features and deep supervision for children with autism spectrum disorder. 2019 Presented at: IEEE International Conference on Multimedia & Expo Workshops; Jul 8-12, 2019; Shanghai. [doi: 10.1109/icmew.2019.00119]
26. Wang S, Jiang M, Duchesne X, Laugeson E, Kennedy D, Adolphs R, et al. Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. Neuron 2015 Nov 04;88(3):604-616 [FREE Full text] [doi: 10.1016/j.neuron.2015.09.042] [Medline: 26593094]
27. Carette R, Elbattah M, Dequen G, Guerin JL, Cilia F. Visualization of eye-tracking patterns in autism spectrum disorder: method and dataset. : IEEE; 2018 Presented at: The 13th International Conference on Digital Information Management; Sep 24-26, 2018; Berlin. [doi: 10.1109/icdim.2018.8846967]
28. Carette R, Elbattah M, Cilia F, Dequen G, Guérin J, Bosche J. Learning to predict autism spectrum disorder based on the visual patterns of eye-tracking scanpaths. 2019 Presented at: The 12th International Joint Conference on Biomedical Engineering Systems and Technologies - HEALTHINF; 2019; Prague. [doi: 10.5220/0007402601030112]
29. Rogé B. Adaptation Française de l'échelle d'évaluation de l'autisme infantile (C.A.R.S.). In: Editions d'Applications Psychotechniques. Issy-les- Moluineaux: Editions d'Applications Psychotechniques; 1989.
30. Guidetti M, Tourrette C. Evaluation de la Communication Sociale Précoce (ECSP). In: Eurotests. Parisurotests Editions. Paris: Eurotests; 2009.
31. Cilia F, Aubry A, Le Driant B, Bourdin B, Vandromme L. Visual exploration of dynamic or static joint attention bids in children with autism syndrome disorder. Front Psychol 2019;10:2187 [FREE Full text] [doi: 10.3389/fpsyg.2019.02187] [Medline: 31649576]
32. Garry C, Cilia F, Landuré M, Aguillon EN, Rovira K, Brisson J. Étude longitudinale de l'orientation sociale chez les enfants avec TSA d'âge préscolaire. Enfance 2017;4:477-481 [FREE Full text] [doi: 10.3917/enf1.174.0477]
33. Cilia F, Aubry A, Bourdin B, Vandromme L. Comment déterminer les zones d'intérêt visuelles sans a priori? Analyse des fixations d'enfants autistes en oculométrie. Revue de Neuropsychologie 2019;11(2):144-150. [doi: 10.1684/nrp.2019.0487]
34. Matplotlib: visualization with Python. Matplotlib. URL: https://matplotlib.org/ [accessed 2020-04-01]
35. Xu Y, Jia R, Mou L, Li G, Chen Y, Lu Y, et al. Improved relation classification by deep recurrent neural networks with data augmentation. 2016 Presented at: The 26th International Conference on Computational Linguistics: Technical Papers; Dec 14, 2016; Osaka.
36. Wang J, Perez L. The effectiveness of data augmentation in image classification using deep learning. Convolutional Neural Networks Vis. Recognit 2017:2017.
37. Chollet F. Keras. Github Repository. URL: https://github.com/fchollet/keras [accessed 2020-04-01]
38. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 May 28;521(7553):436-444. [doi: 10.1038/nature14539] [Medline: 26017442]
39. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 2014;15:1929-1958.
40. Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ, et al. Detecting novel associations in large data sets. Science 2011 Dec 16;334(6062):1518-1524 [FREE Full text] [doi: 10.1126/science.1205438] [Medline: 22174245]
41. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. Nat Rev Genet 2012 Mar 13;13(4):260-270 [FREE Full text] [doi: 10.1038/nrg3182] [Medline: 22411464]
42. Maurice C, Haiser H, Turnbaugh P. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. Cell 2013 Jan 17;152(1-2):39-50 [FREE Full text] [doi: 10.1016/j.cell.2012.10.052] [Medline: 23332745]
43. Albanese D, Filosi M, Visintainer R, Riccadonna S, Jurman G, Furlanello C. Minerva and minepy: a C engine for the MINE suite and its R, Python and MATLAB wrappers. Bioinformatics 2013 Feb 01;29(3):407-408. [doi: 10.1093/bioinformatics/bts707] [Medline: 23242262]
44. Nebout A, Wei W, Liu Z, Huang L, Le MO. Predicting saliency maps for ASD people. 2019 Presented at: International Conference on Multimedia & Expo Workshops; July 8-12, 2019; Shanghai. [doi: 10.1109/icmew.2019.00121]

45. Ju Z, Gun L, Hussain A, Mahmud M, Ieracitano C. A novel approach to shadow boundary detection based on an adaptive direction-tracking filter for brain-machine interface applications. Applied Sciences 2020 Sep 27;10(19):6761. [doi: 10.3390/app10196761]

Abbreviations

ADI-R: Autism Diagnostic Interview-Revised
ADOS-G: Autism Diagnostic Observation Schedule-Generic
ASD: autism spectrum disorder
CARS: Childhood Autism Rating Scale
CNN: convolutional neural network
ESCS: Early Social Communication Scale
MIC: maximal information coefficient
ML: machine learning
POG: point of gaze
ROC: receiver operating characteristic

Edited by A Kushniruk; submitted 03.02.21; peer-reviewed by Y Xiong, C Ieracitano; comments to author 28.03.21; revised version received 29.04.21; accepted 24.05.21; published 25.10.21.

Please cite as:
Cilia F, Carette R, Elbattah M, Dequen G, Guérin JL, Bosche J, Vandromme L, Le Driant B
Computer-Aided Screening of Autism Spectrum Disorder: Eye-Tracking Study Using Data Visualization and Deep Learning
JMIR Hum Factors 2021;8(4):e27706
URL: https://humanfactors.jmir.org/2021/4/e27706
doi:10.2196/27706
PMID:34694238

Original Paper

Perceptions of Patients and Physicians on Teleconsultation at Home for Diabetes Mellitus: Survey Study

Nazaré Rego[1,2], PhD; Helena Silva Pereira[1], MSc; José Crispim[3], PhD

[1]Escola de Economia e Gestão, Universidade do Minho, Braga, Portugal

[2]Institute for Systems and Computer Engineering, Technology, and Science (INESC TEC), Porto, Portugal

[3]Núcleo de Investigação em Políticas Económicas e Empresariais (NIPE), Escola de Economia e Gestão, Universidade do Minho, Braga, Portugal

Corresponding Author:
Nazaré Rego, PhD
Escola de Economia e Gestão
Universidade do Minho
Campus de Gualtar
Braga, 4710 - 057
Portugal
Phone: 351 253604565
Email: nazare@eeg.uminho.pt

Abstract

Background: Diabetes mellitus (DM) is one of the most challenging diseases in the 21st century and is the sixth leading cause of death. Telemedicine has increasingly been implemented in the care of patients with DM. Although teleconsultations at home have shown to be more effective for inducing HbA_{1c} reduction than other telemedicine options, before the 2019 coronavirus disease crisis, their use had been lagging behind. Studies on physicians' or patients' perceptions about telemedicine have been performed independently of each other, and very few have focused on teleconsultations. In a time of great pressure for health systems and when an important portion of health care has to be assured at a distance, obtaining insights about teleconsultations at home from the stakeholders directly involved in the health care interaction is particularly important.

Objective: The perceptions of patients and physicians about their intentions to use home synchronous teleconsultations for DM care are examined to identify drivers and barriers inherent to programs that involve home teleconsultations.

Methods: Two identical questionnaires integrating the technology acceptance model and the unified theory of acceptance and use of technology and assessing the confidence in information and communication technology use of patients and physicians were developed. Responses by patients (n=75) and physicians (n=68) were analyzed using canonical correlation analysis.

Results: Associations between predictor constructs (performance, effort, social influence, facilitating conditions, and attitude) and intention to use yielded significant functions, with a canonical R^2 of 0.95 (for physicians) and 0.98 (patients). The main identified barriers to patient intention to use were the expected effort to explain the medical problem, and privacy and confidentiality issues. The major drivers were the facilitation of contact with the physician, which is beneficial to patient disease management and treatment, time savings, and reciprocity concerning physicians' willingness to perform teleconsultations. Responses from physicians revealed an association between intention to use and the expected performance of home teleconsultations. The major barrier to intention to use expressed in physicians' answers was doubts concerning the quality of patient examination. The major drivers were time savings, productivity increases, improvements in patient's health and patient management, National Health System costs reduction, and reciprocity relative to patients' willingness to engage in teleconsultations.

Conclusions: To promote the use of home teleconsultations for DM, decision makers should improve patients' health literacy so the physician–patient communication is more effective; explore information and communication technology developments to reduce current limitations of non–face-to-face examinations; ensure patient privacy and data confidentiality; and demonstrate the capabilities of home teleconsultations to physicians.

(JMIR Hum Factors 2021;8(4):e27873) doi:10.2196/27873

KEYWORDS

teleconsultation; diabetes mellitus; telemedicine; eHealth; mobile phone

Introduction

Background

Diabetes mellitus (DM) is one of the most challenging diseases in the 21st century. It is the sixth leading cause of death globally [1] and continues to increase in prevalence, with macro-and microvascular complications resulting in increased disability and huge health care costs [2].

In Portugal, there were 591-699 new cases of diabetes per 100,000 inhabitants in 2015, representing an expense of 0.7%-0.9% of the Portuguese gross domestic product, and 8%-10% of the total spending on health [3]. The country has one of the highest age-adjusted prevalence of diagnosed type 1 or 2 diabetes in the population aged 20 to 79 years in Europe (9.8%) [4].

Telemedicine includes remote patient monitoring using devices (eg, mobile apps) to remotely collect and send data to health care providers, asynchronous interactions to transmit diagnostic images, vital signs, or video clips, along with patient data for later review, and synchronous live videoconferencing consultation among patients and physicians (eg, teleconsultations) or among physicians and specialist health services [5]. Information and communication technology (ICT) has been increasingly implemented in the care of people with DM to improve patient outcomes in areas such as blood glucose management, diet, medication, and exercise monitoring [6]. Although (remote) teleconsultations at home (TH) have been found to be more effective in inducing HbA_{1c} reduction when compared with other telemedicine services, such as remote telemonitoring, tele-education, and telecase management, they have been much less adopted than other forms of telemedicine in type 2 DM care [7]. Real-world data show that, before the COVID-19, teleconsultation appointments as a proportion of clinical activity ranged from 2% among a diabetic cohort to 22% among postoperative patients with hepatobiliary cancer [8,9]. In Spain, teleconsultations were used by only 7 (6.9%) of the 102 that used telemedicine in a sample of 1063 patients with type 2 DM, but obtained the highest rate of satisfaction [10]. Studies on teleconsultation in DM are scarce (eg, [11-14]). Given the use level and care potential of synchronous TH, this study investigated the necessary conditions to encourage their use.

Studies on physicians' (eg, [15]) or patients' perceptions (eg, [6,16]) about telemedicine have been performed independently of each other. As both groups are essential to the use of these services, this study surveyed the perceptions of the two using an identical data collection instrument and compared the results of the analysis of their responses. Identical questionnaires for both patients and physicians were used because, according to the literature [17], the factors affecting their willingness to adopt teleconsultations were the same. A canonical correlation analysis (CCA) was performed to find associations among a set of predictor constructs derived from the technology adoption literature and the *intention of use* (*IoU*) of TH by patients with DM and their physicians. CCA is a multivariate statistical technique used to study the interrelationships among sets of multiple dependent and independent variables [18]. It is an appropriate and powerful multivariate technique to identify the underlying independent relationship between the 2 sets of variables of the studied model because of the high number of variables in each construct.

Objective

In summary, this study assesses the perspective of patients with DM and physicians regarding the drivers and barriers inherent to programs that involve patients with DM teleconsulting with their physicians from their homes.

Methods

Research Model

Overview

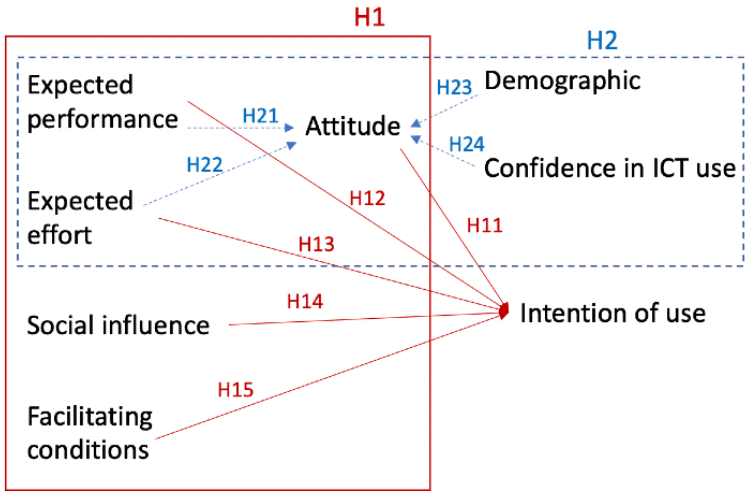
The questionnaires for both populations were based on the integration of the technology acceptance model (TAM) [19,20] and the unified theory of technology of acceptance and use (UTAUT) [21]. The full version of the data collection instruments (in Portuguese) can be seen in Multimedia Appendix 1 [15,17,21-25]; the English translation of the questions can be implied from the row titles of the table in Multimedia Appendix 2 [19,20,26-31].

According to Davis [19], an individual tends to use (or not use) a new technology if they identify an improvement in their professional performance. That is, if he or she easily identifies the *perceived utility*. However, the author states that the usefulness of that technology will only be recognized if the effort to learn how to use it is not very high, that is, if the use of the technology compensates for the learning effort—*perceived ease of use*. The TAM uses these 2 main constructs to influence the actual use of technology. Both have an independent effect on *IoU*, as people form intentions to adopt certain behaviors that can improve their performance at work if the effort required to learn a new technology is not considerable [19-21].

Yarbrough and Smith [32] and Holden and Karsh [26] reviewed articles on the applicability of TAM in health, reaching similar conclusions; the constructs have been repeatedly validated and the variance of the dependent variable *IoU* or *actual use of technology* has been widely explained (between 40% and 70%, depending on the study).

However, TAM is not very sensitive in identifying barriers to the acceptance of technology, which may influence all TAM variables. Thus, new theories explaining the acceptance of technology have emerged. One of these theories is the UTAUT, developed by Venkatesh et al [21]. UTAUT integrates the essential constructs of 8 models of technology and considers 4 constructs that directly influence the intention to use the technology: *expected performance* (P), *expected effort* (E), *social influence* (S), and *facilitating conditions* (F). The research model of this study (Figure 1) integrates TAM with UTAUT, adding an *attitude* construct [15,17,33]. In addition, we tested whether *confidence in ICT use* [27] and *demographic characteristics* were associated with *attitude*. An eventual relationship between gender and *attitude* was explored, and it was hypothesized that a younger age and higher qualifications could favor *attitude* [34].

Figure 1. Research model. ICT: information and communication technology.



Hypotheses

- H1: Do the predictors derived from the literature positively influence the IoU of TH? (suphypotheses in Table 1).
- H2: Do the predictors derived from the literature positively influence attitude (suphypotheses in Table 2).

The analysis of associations among constructs resulted in the identification of major drivers and barriers to DM (synchronous) TH.

Table 1. Subhypotheses of hypothesis 1.

Subhypotheses	Predictor	Effect
H11	Attitude	Positively influences the *intention of use* of teleconsultations at home
H12	Expected performance	Positively influences the *intention of use* of teleconsultations at home
H13	Expected effort	Positively influences the *intention of use* of teleconsultations at home
H14	Social influence	Positively influences the *intention of use* of teleconsultations at home
H15	Facilitating conditions	Positively influences the *intention of use* of teleconsultations at home

Table 2. Subhypotheses of hypothesis 2.

Subhypotheses	Predictor	Effect
H21	Expected performance	Positively influences *attitude*
H22	Expected effort (ie, perceived ease of use [20])	Positively influences *attitude*
H23	Demographic characteristics	Positively influences *attitude*
H24	Confidence in information and communication technology use	Positively influences *attitude*

Sample and Scales

Data were collected from patients with type 1 or 2 diabetes or their caregivers, in case of child patients (75 valid responses) and physicians (68 valid responses) selected by rational choice and snowball sampling (as highly specific populations were at stake) from the north of Portugal during the fourth quarter of 2018. Concerning the patients, 51 questionnaire answers (51/75, 68% of total valid answers) were collected in-person and in paper at primary care centers belonging to the Group of Primary Care Centres of Braga, an organization that coordinates 22 primary care centers; the other were collected on web through DM patients' associations. For physicians, the answers were collected on web with the collaboration of the same group of primary care centers. This organization sent an email with a link to the questionnaire to their physicians.

The perceptions of both groups were measured using 2 identical questionnaires based on a 7-point concordance Likert scale and a 5-point confidence Likert scale (Multimedia Appendix 2).

Table 3. Characteristics of the samples.

Characteristics	Patients (or caregivers; n=75) n (%)	Physicians (n=68), n (%)
Type of respondent		N/A[a]
Patient	61 (81)	
Caregiver	14 (19)	
Gender		
Female	38 (51)	47 (69)
Male	37 (49)	21 (31)
Education		N/A
Basic or less	33 (44)	
Secondary	16 (21)	
Bachelor	8 (21)	
Master	7 (9)	
Opted to not respond	11 (15)	
Medical specialty	N/A	
General practitioner		52 (77)
Other		16 (23)
Financial situation (ability to live with monthly budget)		N/A
Faces difficulties	8 (11)	
Needs to manage carefully	25 (33)	
Can go through	25 (33)	
Goes through easily	14 (19)	
Goes through very easily	3 (4)	
DM[b] type		N/A
1	29 (39)	
2	44 (59)	
Other	2 (3)	
Treatments or disease control		N/A
Oral antihyperglycemic	42 (56)	
Insulin	36 (48)	
Antihypertensive	35 (73)	
Antidyslipidemia	31 (41)	
Physical exercise	41 (55)	
Diet	44 (59)	
Daily auto monitoring of the disease	35 (73)	
Local for DM consultations		
Primary care center (public)	52 (69)	52 (77)
Public hospital	34 (45)	21 (31)
Private hospital	8 (11)	8 (12)
Other (private)	3 (4)	3 (4)
Mode of transport to consultations		N/A
By car	50 (67)	
By bus	21 (28)	

Characteristics	Patients (or caregivers; n=75) n (%)	Physicians (n=68), n (%)
On foot	19 (26)	
Other	3 (4)	
Electronic devices use		
Computer	28 (37)	60 (88)
Laptop	35 (47)	30 (44)
Tablet	17 (23)	15 (22)
Smartphone	67 (90)	49 (72)
None	6 (8)	0 (0)
Use of app for real-time video call		
Never used	28 (37)	6 (9)
Rarely	15 (20)	26 (38)
Once per month	7 (9)	12 (18)
Once per week	7 (9)	5 (7)
Several times in a week	7 (9)	9 (13)
Everyday	11 (15)	10 (15)

[a]N/A: not applicable.

[b]DM: diabetes mellitus.

Data Analysis

CCA was used to analyze the correlation between the set of dependent variables (*IoU* construct) and the set of predictor constructs. This method is useful when variables have multiple causes and effects, similar to the complex reality of human behavior and cognition. The computations were performed using SPSS (version 24.0; IBM Corp).

The average of the observed values is often used to form the constructs with consequent smoothing of the responses, which can lead to constructs that do not contain the variability expressed in the measurement indicators. CCA examines the relationship between the 2 observed variable sets without having this disadvantage.

Variables with a canonical correlation of 0.45 or above were considered in the final CCA model. The reliability statistics measured by Cronbach α for each construct scale were very good for *expected performance* (.87 for physicians and .83 for patients), *facilitating conditions* (.77 for physicians and .74 for patients), *attitude* (.83 for physicians and .82 for patients), and *IoU* (.91 for physicians and .78 for patients), and acceptable for *expected effort* (.60 for physicians and .56 for patients), and *social influence* (.62 for physicians and .56 for patients).

Results

Sample Characteristics

Survey responses of 75 patients (Table 3 and Multimedia Appendix 2) aged 10-86 years (mean 51, SD 17.1) were obtained. Of the 75 respondents, 6 (8%) had never used computers, smartphones, or tablets (Table 3). The average age of these 6 patients was 73 (SD 8.9, range 61-84) years. On average, patients had 3.1 DM consultations per year (range: 1-12). According to the respondents, DM consultations took 133 minutes on average (including travel and waiting). Of the 75 respondents, 33 (44%) patients or caregivers felt very or extremely confident and 12 (16%), moderately confident using computers or the internet. Moreover, 40% (30/75) were very or extremely confident and 24% (18/75) were moderately confident in the use of real-time video call apps. More than half had heard about telemedicine (43/75, 57% of patients) and 30% (23/75) about teleconsultations, but only 2 had participated in one in real time, and 37% (28/75) had never used an app to make a real-time video call.

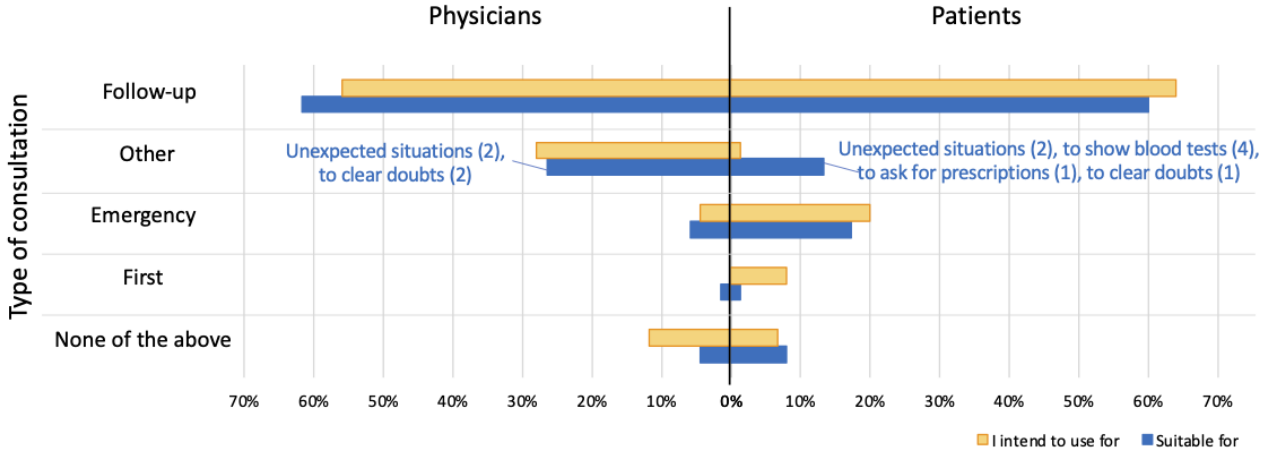
In total, 68 valid responses from physicians aged 25-63 years (47/68, 69% of physicians in the interval 26-35 years) were received. Of the 68 respondents, 46 (67%) performed between 10 and 40 consultations per month, with an average duration of 23 minutes. Only 6 (9%) out of 68 physicians had never used a video call app. Moreover, 81% (55/68) felt very or extremely confident and 19% (13/68) moderately confident using computers or the internet, and 54% (37/68) were very or extremely confident. Furthermore, 28% (19/68) were moderately confident in the use of real-time video call apps. Of the 68 respondents, 33 (48.5%) physicians had never heard of TH, and 8 (12%) had already carried out synchronous teleconsultations. Although 56% (38/68) of physicians stated that they intended to use TH in follow-up consultations, 34% (25/68) answered that they would not use TH because they did not consider them a good method for health provision.

The distribution of concordance scores showed a significant variability in both groups. Wilcoxon Mann-Whitney tests identified differences between physicians and patients' responses. Physicians had higher confidence in ICT use, but they also had higher scores for item E_3—*Will only use TH if easy to learn* and S_1—*if there was technical assistance*. Patients

had (1) in general, a more favorable *attitude* toward TH, and higher scores in the perception that (2) *TH can invade their privacy* (F_3), but (3) *be faster* (P_1), (4) *the medical problem can be correctly understood* (E_1) in a TH, and (5) *will have TH whenever the counterpart wants to* (S_3).

Both patients and physicians considered follow-up to be the best purpose for TH (Figure 2).

Figure 2. Intention of use or suitability by type of consultation.



Hypotheses Testing

H1: Predictors Positively Influence the IoU of TH

Overview

For both physicians and patients, at least one variable of the predictors is associated with *IoU* variables (H11, H12, H13, H14, and H15 cannot be rejected).

Figure 3 shows the association between the latent variable *IoU* and the related covariate set of variables (predictors). Tables 4 and 5 present a validation, through comparison with the literature, of the revealed associations.

Figure 3. Canonical associations between predictors and intention of use. CCA: canonical correlation analysis.

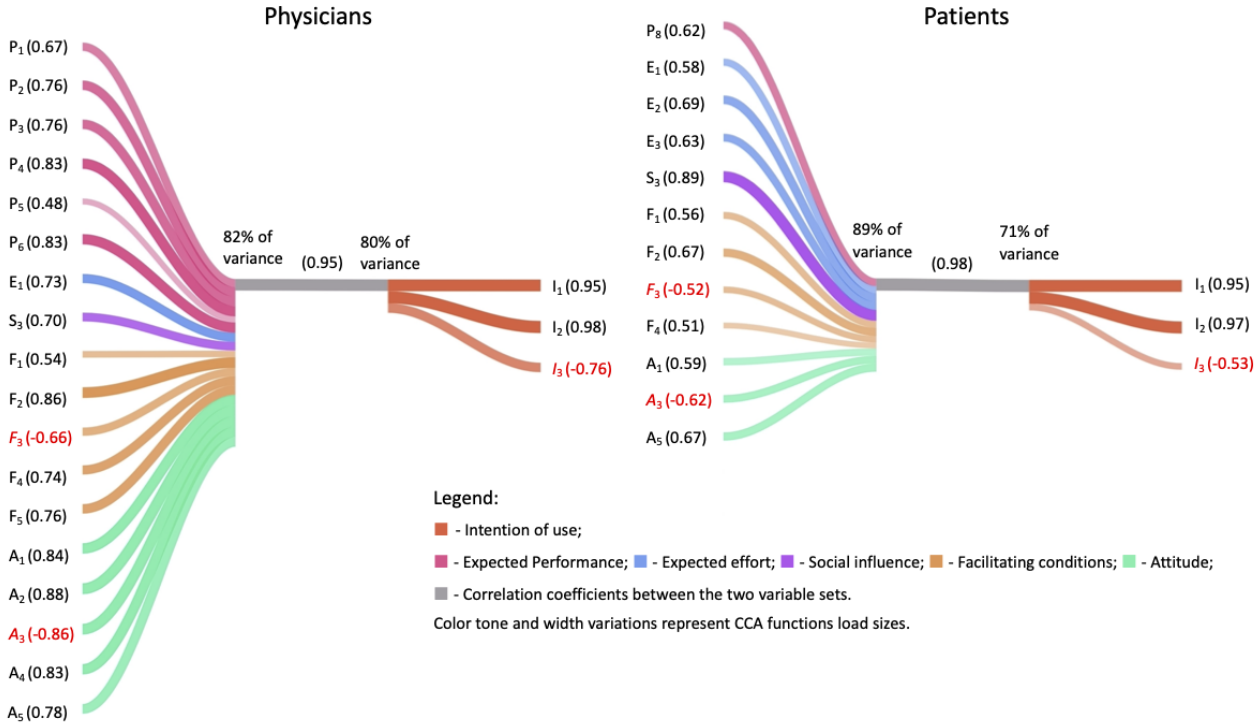


Table 4. Validation of the revealed associations for patients.

Canonical correlation analysis association	Variable	Literature
Primary contributors		
Expected effort—IoU[a]	E_2 (I can explain my medical problems using the computer)	Several studies found that the medium allowed patients to *open up* more than face-to-face consultations and that they felt empowered to ask more questions [35,36]
Social influence—IoU	S_3 (I'll do teleconsultations whenever the physician wants to)	To boost use, physician support and recommendation is necessary [10]
Facilitating conditions—IoU	F_2 (beneficial in my management of my disease)	In Spain, most patients with type 2 diabetes (73.6%) considered that the use of telemedicine had optimized (quite a bit or a lot) the management of their disease [10]
Attitude—IoU	A_7 (will not increase the provision of health care services)	Several studies found that patients were satisfied with teleconsultations, but also that they would still want the option to attend in person as they believe it to be the *gold standard* [36,37]
Secondary contributors		
Expected performance—IoU	P_8 (allows me to save my time)	Waiting times were shorter for patients seen by teleconsultation than in face-to-face consultation as they bypassed the normal admission processes [38]
Effort expectancy—IoU		
	E_1 (physician can correctly understand my medical problem)	Physical examination has become a ritual, expected, and performed as tradition rather than clinical usefulness [39]. For the time being, teleconsultations in outpatient settings are most likely to be confine to dialogue-based consultations where the need for rigorous physical examination is absent [8]
	E_3 (perceived as being easy to learn)	The patients were very satisfied with the technology, no major problems with its use; nearly 100% of patients reported that they would use it again and recommend it [40]
Facilitating conditions—IoU		
	F_1 (can facilitate contact with the physician)	Several studies found that improved access to care was associated with patient satisfaction [40]
	F_3 (can invade patient's privacy) [b]	In a study of a teenaged population, parents are worried that the connection might not be secure enough to ensure privacy and patients fear that they might be overheard by family [36]
	F_4 (will not interfere with confidentiality of my health data)	The need to ensure the security and confidentiality of patient records diminishes the preference for and use of telemedicine technology [41]
Attitude—IoU		
	A_1 (it is a good way to provide health care services)	In the United Kingdom, teleconsultations for acute stroke management had item values (like morbidity, mortality, and discharge rates) comparable with national standards [42,43]
	A_3 (it is unpleasant to use teleconsultations at home)	Several studies found that patients were satisfied with teleconsultations but also that they would still want the option to attend in person as they believe it to be the *gold standard* [36,37]
	A_5 (teleconsultations at home can be a supplemental health care service)	Several studies found that patients were satisfied with teleconsultations but also that they would still want the option to attend in person as they believe it to be the *gold standard* [36,37]

[a]IoU: intention of use.

[b]Variables in *italic* had a negative sign in the predictors set.

Table 5. Validation of the revealed associations for physicians.

Canonical correlation analysis association	Variable	Literature
Primary contributors		
Expected performance—IoU[a]		
	P_2 (improve my productivity)	Benger et al [38] refer that teleconsultations are as much as 4 times as long as their face-to-face equivalent; however more recent studies, found them to be shorter in length (eg, [9])
	P_3 (improves management of patient care)	Workload can be classified as the biggest workflow-related concern, as it was overrepresented in the results, being addressed in 12 of the 23 studies analyzed in the systematic literature review by Granja et al [44]
	P_4 (improves the patient's health)	Telehealth is a safe option for delivery of self-management support [45]
	P_6 (improve the effectiveness of my work)	Several examples of real-world evaluations of working teleconsultation services have demonstrated that they can achieve meaningful reductions in did not attend (DNA) rates [46]
Social influence—IoU	S_3 (I'll do teleconsultations whenever the patient wants to)	The literature emphasizes the role of physicians in promoting telemedicine use [10]
Facilitating conditions—IoU		
	F_2 (beneficial in my patient management and treatment)	DNA rates were lower (13% vs 28%) and HbA_{1c} control improved in patients that chose to attend by teleconsultation [47]
	F_4 (will not interfere with confidentiality of the patient's health data)	Lack of policies that guarantee the patient's privacy and confidentiality when using and transferring information, lack of authentication by health professionals, and lack of attribution of responsibility for the quality of services are barriers to the adoption of telemedicine in health services [48]
	F_5 (may reduce the costs of the National Health System)	In the past, the use of telemedicine was strongly dependent of technology costs (eg, [49]). Nowadays, technology allows cost savings: a report on telehealth services in Scotland found that teleconsultations for a 10-week rehabilitation course could be delivered for 3% to 10% of the cost associated with an outreach model (in which the therapist travels) or a centralized model (in which the patient travels), with savings primarily being delivered through reduced travel costs [50]
Attitude—IoU		
	A_1 (good way of providing health care services)	O'Cathail et al [8] summarize contradictory studies: some show that physicians lack confidence in their teleconsultation diagnosis; others assessed the concordance of diagnosis in both an inpatient and outpatient setting in neurology and found 96%-100% of cases were accurately diagnosed and managed via teleconsultation
	A_2 (it is a good idea to use teleconsultations at home)	Amid the COVID-19 pandemic: patients can keep in touch with their routine physicians via teleconsultations; physicians could ensure drug compliance; educate patients and their caregivers; make patients aware of the common symptoms of hypoglycemia; and help patients cope with psychological problems [51]
	A_3 (it is unpleasant to use teleconsultations at home) [b]	The opening phase of the consultation was found to be unfamiliar, leading to interruptions and apologies on both sides whereas a dialogue flow was established [52]
	A_4 (teleconsultation will be a common method in the future)	*No literature exploring this specific variable association was found*
	A_5 (teleconsultations at home can be a supplemental health care service)	In some cases, the inability to perform some aspects of physical examination is likely to restrict video outpatient teleconsultations utility for more *routine* outpatient appointments [9]
Secondary contributors		
Expected performance—IoU		

Canonical correlation analysis association	Variable	Literature
	P_1 (I will be able to complete the patient's medical consultation more quickly)	According to Benger et al [38], teleconsultations were, on average, almost twice as long as their face-to-face equivalent. However, more recent studies found them to be shorter (eg, [9])
	P_5 (I will be able to examine the patient as well as I would during face-to-face consultations)	O'Cathail et al [8] review shows that: (1) a lack of confidence on teleconsultation diagnosis exists among professionals, (2) studies in neurology assessed the concordance of diagnosis in both an inpatient and outpatient setting and found that 96%-100% of cases were accurately diagnosed and managed via teleconsultation
Expected effort—IoU	E_1 (I can understand the medical problem correctly)	O'Cathail et al [8] review shows that: (1) a lack of confidence on teleconsultation diagnosis exists among professionals, (2) studies in neurology assessed the concordance of diagnosis in both an inpatient and outpatient setting and found that 96%-100% of cases were accurately diagnosed and managed via teleconsultation
Facilitating conditions—IoU		
	F_1 (teleconsultations at home facilitate contact with the patient)	Morris et al [47] show that, among a diabetic cohort, teleconsultation improved the DNA rate from 28% to 13% and HbA_{1c} control
	F_3 (teleconsultations at home can invade patient's privacy)	Some health professionals thought teleconsultations were an invasion of patients' personal space [36]

[a]IoU: intention of use.

[b]Variables in *italic* had a negative sign in the predictors set.

Patients

The analysis between the set of predictor variables and *IoU* yielded one significant function with a canonical correlation of 0.98 (*P*<.*001*) and a canonical R^2 of 0.93. The model explains about 98% (1–Wilk λ=1–0.01789) of the variance shared among the variable sets.

E_2 (can explain medical problems using a computer), S_3 (will have TH whenever the counterpart wants to), F_2 (will be beneficial to manage the disease), and A_5 (can be a supplemental care service) were the primary contributors to the predictor synthetic variable.

P_8 (will save time), E_1 (medical problem can be correctly understood), E_3 (will only be used if easy to learn), F_1 (facilitates contact with counterpart), F_3 (can invade privacy), F_4 (will not interfere with confidentiality of health data), A_1 (is a good way to provide health care), and A_3 (will be unpleasant to use TH to receive health care) were secondary contributors. The coefficient of I_3 (will not be used routinely) is negative because it was negatively related to all the predictors except F_3 and A_3: the perception that technology *can invade patients' privacy* (F_3) and that *to use teleconsultation will be unpleasant* (A_3) were positively associated with *not using teleconsultation routinely* (I_3). These results generally support the theoretically expected relationships (Table 4).

Physicians

The analysis between the predictors and *IoU* yielded one significant function with a canonical correlation of 0.95 (*P*<.*001*) and a canonical R^2 of 0.91. The model explains about 99% (1–Wilk λ=1–0.01010) of the variance shared between the variable sets.

The primary contributors to the predictor synthetic variable were TH can improve *my productivity* (P_2), *management of patient care* (P_3), *the patient's health* (P_4), *the effectiveness of my work* (P_6), E_1 (medical problem can be correctly understood), S_3 (will have TH whenever the counterpart wants to), F_2 (will be beneficial to manage patients and their treatment), F_4 (will not interfere with confidentiality of health data), F_5 (can decrease the National Health System costs), A_1 (is a good way to provide health care services), A_2 (is a good idea to use TH), A_3 (will be unpleasant to use TH to provide health care), A_4 (a common method for providing health care in the future), and A_5 (can be a supplemental health care service). P_1 (medical consultation can be completed faster), P_5 (patient examination is as good as in face-to-face consultations), F_1 (facilitates contact with counterpart), and F_3 (can invade patient's privacy) were secondary contributors. The structure coefficient of I_3 is negative; therefore, F_3 and A_3 are positively associated with I_3. All other significant predictors were negatively associated with the I_3. The results are described in Table 5.

Perceptions of Patients Versus Perceptions of Physicians

In terms of *expected performance* (H12), for physicians, all the variables were statistically associated with *IoU*. For the patients, only P_8 (economy of time) was statistically associated with *IoU*, but the loading was relatively low (0.62). On the contrary, for the patients, all *expected effort* (H13) variables were statistically associated with *IoU*, whereas for physicians, only E_1 (being able to understand the medical problem correctly).

In terms of *social influence* (H14), the theoretical relationships from the literature were not confirmed: only S_3 (willingness to do TH whenever the counterpart wants to) exhibited statistically significant associations with *IoU*.

Relative to *facilitating conditions* (H15) and *attitude* (H11), the results for both groups were generally supportive of the theoretically expected relationships (with higher loadings for physicians). All facilitating conditions except F_5 (TH can decrease National Health System costs) were statistically associated with *IoU* for the 2 groups. F_5 was statistically significant only for the physicians. In terms of *attitude*, the perception that TH was *a good way to provide health care services* (A_1) *as supplemental care* (A_5) was positively associated, and *unpleasant to use to provide or receive health care* (A_3) was negatively associated with *IoU* for both groups. Perceiving that TH will be *a common method in the future* (A_4) was positively associated with *IoU* only for physicians.

H2: The Predictors Derived From the Literature Positively Influence Attitude

Demographic data were not associated with *attitude* and, for physicians, confidence in ICT use. H21 and H22 cannot be rejected for both groups, and H24 (specifically, relative to C3—*confidence in making videocalls*) cannot be rejected for patients.

Figure 4 shows the canonical associations between the predictors and *attitude*. For physicians (canonical R^2 of 0.96), *expected performance* and E_1 (understanding the patient's medical problem correctly) were associated with *attitude*, all as primary contributors. For the patients (canonical R^2 of 0.85), higher associations were observed with *expected effort*, C_3, P_1 (medical consultation can be completed faster), P_4 (can improve patient's health), and P_8 (will save patient's time). P_1, P_8, and E_1 are the primary contributors. Table 6 synthetizes the major drivers and barriers of TH for the DM.

Figure 4. Canonical associations between predictors and attitude. CCA: canonical correlation analysis.

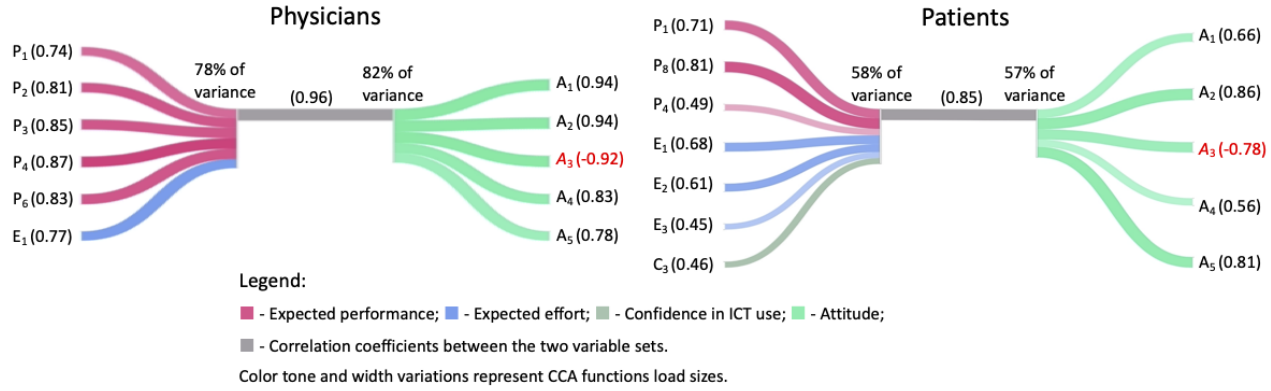


Table 6. Home teleconsultation barriers and drivers.

Categories used to predict intention to use and attitude			
Barriers		Drivers	
Patients	Physicians	Patients	Physicians
Expected performance			
• None identified	• P_5—able or unable to examine the patient as well as he or she would in face-to-face consultations	• P_8—saves patient's time	• P_1—consultation will be faster • P_2—improves physician's productivity • P_3—improves patient management care • P_4—improves patient's health • P_6—improves effectiveness of physician's work
Expected effort			
• E_1- physician can (not) understand patient's medical problem correctly • E_2—patient can explain her/his medical problems using the computer • E_3—patient will only use if it is easy to learn	• E_1- physician can (not) understand patient's medical problem correctly	• None identified	• None identified
Social influence			
• None identified	• None identified	• S_3—willingness to do TH[a] whenever the physician or patient wants to	• S_3—willingness to do TH whenever the physician or patient wants to
Facilitating conditions			
• *F_3—can invade patient's privacy*[b] • F_4—use interferes with confidentiality of patient's health data	• *F_3—can invade patient's privacy* • F_4—use interferes with confidentiality of patient's health data	• F_1—facilitates contact with the patient or physician • F_2—beneficial to patient management and treatment	• F_1—facilitates contact with the patient or physician • F_2—beneficial to patient management and treatment • F_5—may reduce the costs of the National Health System
Attitude			
• *A_3—it is unpleasant for physician–patient relationship* • A_5—should (only) be a supplemental health care service	• *A_3—it is unpleasant for physician–patient relationship* • A_5—should (only) be a supplemental health care service	• A_1—is a good way of providing health care services	• A_1—is a good way of providing health care services • A_2—it is a good idea to provide TH • A_4—TH will be a common method in the future
Confidence in Information and communication technology use			
• None identified	• None identified	• C_3—confidence in making videocalls	• C_3—confidence in making videocalls

[a]TH: teleconsultations at home.

[b]Variables in italic had a negative sign in the predictors set.

Discussion

Principal Findings

The main contribution of this study is the identification of relationships among a set of construct predictors taken from the literature and the intention to use (synchronous) home teleconsultations. Obtaining insights about home teleconsultations from the stakeholders directly involved in the health care interaction—that is, patients and physicians, is particularly important in a time of great pressure for health systems and when an important portion of health care has to be assured at a distance.

TH appear to be safe and effective in appropriate clinical situations [9]. In addition, it should not be forgotten that more vulnerable fringes of the population would not have the resources needed for this type of consultation; for example, in our sample, some older patients had never used computers, smartphones, or tablets. Physicians and patients will likely be supportive of their use if they are offered as supplemental and in support of traditional care models rather than to replace them (most of the 2 samples agreed with this type of teleconsultation; Multimedia Appendix 2, variable A_5). This result is in line with the findings of Gilbert et al [37], from the patient's perspective, and Greenhalgh et al [9], from that of the physicians.

Health illiteracy and the physical examination *ritual* (referred by Haig-Ferguson et al [36] in the context of a pediatric chronic fatigue service) may explain why patients see TH only as an extra health care service. Patients should be encouraged and supported for their use.

Expected performance factors (time savings, increased productivity or efficiency, better disease management, health improvement, and quality of the clinical examination) were the most important factors for *intention to use* among physicians, which is in line with the literature (Table 5). On the contrary, except for time savings, patients' perceptions did not reveal an association between performance variables and *intention to use*.

Another difference concerns the *expected effort* needed to use TH. For the patients, explaining *and being understood when communicating their medical problems using the computer* and *the technology being easy to learn* are positively associated with *intention to use.* This type of concern has been identified in the literature [8]. For physicians, these factors are not related to *the intention to use*, except for the necessary effort to *understand the patient's problem*. In their systematic literature review on physicians' eHealth adoption, Granja et al [44] concluded that the major facilitators of eHealth are the quality of the diagnosis and patient-centered care.

For both groups, the only *social influence* variable associated with *intention to use* was the willingness of patients or physicians to participate in TH if their counterpart wants to, which means that each group can encourage the use of the other. The literature has only referred to the importance of physicians' recommendations for teleconsultation [10].

For both patients and physicians, all *facilitating conditions* and *attitude* variables toward using TH were associated with *intention to use*, which is in accordance with the literature (Tables 4 and 5).

Curiously, contrary to the evidence described in the literature (eg, [53]), this study did not find a statistically significant association between *demographic data* or *confidence in ICT use* and *attitude* for any of the groups, with the exception of confidence in using video calls for patients. The current COVID-19 pandemic has led many people to communicate through videoconferencing. Given that our results point to a positive association between confidence in the use of video calls and *attitude*, the pandemic situation may have been a booster for patients' TH adherence. For teleconsultations to be an effective addition to health services beyond COVID-19 they should be considered not only as a technological issue, but also as a complex organizational change problem [54]; from the aspects raised by the authors, we would like to highlight the need for adjusted legal frameworks and reimbursement schemes.

Major barriers to TH use identified were: (1) the inability to correctly understand the medical problem, (2) threats to patient privacy, (3) health data confidentiality, (4) unpleasantness of TH to provide or receive health care, and (5) type of TH use (supplemental care service). On the basis of the perceptions of patients, costs do not seem to be a barrier to TH use, contrary to what has been described in the literature [8]. Probably, this result was observed because, nowadays, technological devices that can be used to make teleconsultations (smartphones, tablets, laptops, etc) are easily available in most of the situations. As digital interaction generally has insignificant costs for patients, cost is highly dependent on the existence of the technology.

The major identified TH drivers were (1) the perception that they facilitate contact, and (2) the fact that the use by each group was highly influenced by the other. Furthermore, physicians are very sensitive to issues related to the performance and quality of service.

The sampling methods limit the generalizability of the results. The composition of the patients' sample in terms of age and education was similar to that of the general population in northern Portugal. However, the proportion of patients with type 1 (type 2) DM in the sample is higher (lower) than expected in the population [4]. Thus, the patients' population may be, on average, older and less educated than the sample in this study. As the data concerning the perceptions of physicians were collected on web, the sample may be, on average, more technology favorable than the population. Nevertheless, both samples included individuals ranging from more positive to more skeptic about TH. In addition, the results were compared with and discussed against the findings of related studies.

A CCA revealed a strong association between the predictors and the set of dependent variables, in line with the literature. The data analysis included a joint critical comparison of the perceptions of patients and physicians. To promote the use of home teleconsultations for DM, decision makers should: (1) improve patient health literacy, as the inability to explain medical problems correctly emerges as a barrier to teleconsultation use; (2) explore ICT developments to reduce current limitations of non–face-to-face examination; (3) ensure patient privacy and data confidentiality; and (4) demonstrate

the capabilities of home teleconsultations to physicians, namely, in terms of the ability to enhance patient–physician communication and to educate patients and their caregivers toward a better management of the disease.

Conclusions

In the future, it would be interesting that research about teleconsultations acceptance incorporated sustainability related aspects like, for example, fuel consumption, carbon emissions, and loss of work productivity. A recent review [55] concluded that, as patients, health care organizations, and nations continue to look toward video consultations as an alternative, it is essential to continue to theorize in this domain.

Acknowledgments

The research work carried out by the second author was supported by an APES (Portuguese Association of Health Economics)—Medtronic scholarship.

The publication of the study will be financed by the National Funds of the FCT (Portuguese Foundation for Science and Technology) within project number UIDB/03182/2020.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaires to collect perceptions of diabetes mellitus patients and physicians about teleconsultations at home.
[PDF File (Adobe PDF File), 365 KB - humanfactors_v8i4e27873_app1.pdf]

Multimedia Appendix 2
Constructs.
[PDF File (Adobe PDF File), 199 KB - humanfactors_v8i4e27873_app2.pdf]

References

1. World Health Organization. Global Report on Diabetes. Geneva: World Health Organization; 2016.
2. Tamayo T, Rosenbauer J, Wild SH, Spijkerman AM, Baan C, Forouhi NG, et al. Diabetes in Europe: an update. Diabetes Res Clin Pract 2014 Feb;103(2):206-217. [doi: 10.1016/j.diabres.2013.11.007] [Medline: 24300019]
3. Diabetes: Factos e Números – O Ano de 2015 − Relatório anual do observatório Nacional da Diabetes – Parte I. Revista Portuguesa de Diabetes. 2017. URL: http://www.revportdiabetes.com/wp-content/uploads/2017/11/RPD-Vol-12-n%C2%BA-1-Mar%C3%A7o-2017-Relat%C3%B3rios-p%C3%A1gs-40-48.pdf [accessed 2021-10-22]
4. Williams R, Colagiuri S, Chan J, Gregg E, Ke C, Lim LL, et al. IDF Atlas, 9th Edition. Brussels, Belgium: International Diabetes Federation; 2019.
5. Gogia S. Fundamentals of Telemedicine and Telehealth. Cambridge, Massachusetts, United States: Academic Press; 2019:1-412.
6. Macdonald EM, Perrin BM, Kingsley MI. Enablers and barriers to using two-way information technology in the management of adults with diabetes: a descriptive systematic review. J Telemed Telecare 2018 Jun;24(5):319-340. [doi: 10.1177/1357633X17699990] [Medline: 28347218]
7. Lee SW, Chan CK, Chua SS, Chaiyakunapruk N. Comparative effectiveness of telemedicine strategies on type 2 diabetes management: a systematic review and network meta-analysis. Sci Rep 2017 Oct 04;7(1):12680. [doi: 10.1038/s41598-017-12987-z] [Medline: 28978949]
8. O'Cathail M, Sivanandan MA, Diver C, Patel P, Christian J. The use of patient-facing teleconsultations in the National Health Service: scoping review. JMIR Med Inform 2020 Mar 16;8(3):e15380 [FREE Full text] [doi: 10.2196/15380] [Medline: 32175911]
9. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. J Med Internet Res 2018 Dec 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
10. Rodríguez-Fortúnez P, Franch-Nadal J, Fornos-Pérez JA, Martínez-Martínez F, de Paz HD, Orera-Peña ML. Cross-sectional study about the use of telemedicine for type 2 diabetes mellitus management in Spain: patient's perspective. The EnREDa2 Study. BMJ Open 2019 Jun 22;9(6):e028467 [FREE Full text] [doi: 10.1136/bmjopen-2018-028467] [Medline: 31230025]
11. Sood A, Watts SA, Johnson JK, Hirth S, Aron DC. Telemedicine consultation for patients with diabetes mellitus: a cluster randomised controlled trial. J Telemed Telecare 2018 Jul;24(6):385-391. [doi: 10.1177/1357633X17704346] [Medline: 28406066]

12. Bertuzzi F, Stefani I, Rivolta B, Pintaudi B, Meneghini E, Luzi L, et al. Teleconsultation in type 1 diabetes mellitus (TELEDIABE). Acta Diabetol 2018 Feb;55(2):185-192. [doi: 10.1007/s00592-017-1084-9] [Medline: 29209814]
13. Fatehi F, Gray LC, Russell AW, Paul SK. Validity study of video teleconsultation for the management of diabetes: a pilot randomized controlled trial. Diabetes Technol Ther 2015 Oct;17(10):717-725. [doi: 10.1089/dia.2014.0416] [Medline: 26181716]
14. Fatehi F, Martin-Khan M, Smith AC, Russell AW, Gray LC. Patient satisfaction with video teleconsultation in a virtual diabetes outreach clinic. Diabetes Technol Ther 2015 Jan;17(1):43-48. [doi: 10.1089/dia.2014.0159] [Medline: 25296189]
15. Rho MJ, Choi IY, Lee J. Predictive factors of telemedicine service acceptance and behavioral intention of physicians. Int J Med Inform 2014 Aug;83(8):559-571. [doi: 10.1016/j.ijmedinf.2014.05.005] [Medline: 24961820]
16. Zobair KM, Sanzogni L, Sandhu K. Expectations of telemedicine health service adoption in rural Bangladesh. Soc Sci Med 2019 Oct;238:112485. [doi: 10.1016/j.socscimed.2019.112485] [Medline: 31476664]
17. Kohnke A, Cole ML, Bush R. Incorporating UTAUT predictors for understanding home care patients' and clinician's acceptance of healthcare telemedicine equipment. J Technol Manag Innov 2014 Jul;9(2):29-41. [doi: 10.4067/S0718-27242014000200003]
18. Hair Jr JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis. Delhi: Cengage Learning; 2018:1-832.
19. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
20. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
21. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
22. Demiris G, Speedie S, Finkelstein S. A questionnaire for the assessment of patients' impressions of the risks and benefits of home telecare. J Telemed Telecare 2000;6(5):278-284. [doi: 10.1258/1357633001935914] [Medline: 11070589]
23. Kifle M, Payton FC, Mbarika V, Meso P. Transfer and adoption of advanced information technology solutions in resource-poor environments: the case of telemedicine systems adoption in Ethiopia. Telemed J E Health 2010 Apr;16(3):327-343. [doi: 10.1089/tmj.2009.0008] [Medline: 20406120]
24. Ministério da Saúde. Despacho 3571/2013. Diário da República n.º 46/2013. 2013. URL: https://dre.pt/dre/detalhe/despacho/3571-2013-1759945 [accessed 2021-11-08]
25. A health telematics policy. World Health Organization - Report of the WHO Group Consultation on Health Telematics. 1998. URL: https://apps.who.int/iris/handle/10665/63857 [accessed 2021-11-08]
26. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
27. Lam MK, Nguyen M, Lowe R, Nagarajan SV, Lincoln M. "I can do it": does confidence and perceived ability in learning new ICT skills predict pre-service health professionals' attitude towards engaging in e-healthcare? Stud Health Technol Inform 2014;204:60-66. [Medline: 25087528]
28. Li J, Talaei-Khoei A, Seale H, Ray P, Macintyre CR. Health care provider adoption of ehealth: systematic literature review. Interact J Med Res 2013;2(1):e7 [FREE Full text] [doi: 10.2196/ijmr.2468] [Medline: 23608679]
29. Dansky KH, Gamm LD, Vasey JJ, Barsukiewicz CK. Electronic medical records: are physicians ready? J Healthc Manag 1999;44(6):440-454. [Medline: 10662431]
30. Kifle M, Mbarika VW, Datta P. Interplay of cost and adoption of tele-medicine in Sub-Saharan Africa: the case of tele-cardiology in Ethiopia. Inf Syst Front 2006 Jul;8(3):211-223. [doi: 10.1007/s10796-006-8780-2]
31. Taylor S, Todd PA. Understanding information technology usage: a test of competing models. Inform Syst Res 1995 Jun;6(2):144-176. [doi: 10.1287/isre.6.2.144]
32. Yarbrough AK, Smith TB. Technology acceptance among physicians: a new take on TAM. Med Care Res Rev 2007 Dec;64(6):650-672. [doi: 10.1177/1077558707305942] [Medline: 17717378]
33. Kim S, Lee K, Hwang H, Yoo S. Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. BMC Med Inform Decis Mak 2016 Jan 30;16:12 [FREE Full text] [doi: 10.1186/s12911-016-0249-8] [Medline: 26831123]
34. Hoogenbosch B, Postma J, de Man-van Ginkel JM, Tiemessen NA, van Delden JJ, van Os-Medendorp H. Use and the users of a patient portal: cross-sectional study. J Med Internet Res 2018 Sep 17;20(9):e262 [FREE Full text] [doi: 10.2196/jmir.9418] [Medline: 30224334]
35. Tachakra S, Rajani R. Social presence in telemedicine. J Telemed Telecare 2002;8(4):226-230. [doi: 10.1258/135763302320272202] [Medline: 12217106]
36. Haig-Ferguson A, Loades M, Whittle C, Read R, Higson-Sweeney N, Beasant L, et al. "It's not one size fits all"; the use of videoconferencing for delivering therapy in a Specialist Paediatric Chronic Fatigue Service. Internet Interv 2019 Mar;15:43-51 [FREE Full text] [doi: 10.1016/j.invent.2018.12.003] [Medline: 30619719]
37. Gilbert AW, Jaggi A, May CR. What is the patient acceptability of real time 1:1 videoconferencing in an orthopaedics setting? A systematic review. Physiotherapy 2018 Jun;104(2):178-186. [doi: 10.1016/j.physio.2017.11.217] [Medline: 29361298]

38. Benger JR, Noble SM, Coast J, Kendall JM. The safety and effectiveness of minor injuries telemedicine. Emerg Med J 2004 Jul;21(4):438-445 [FREE Full text] [Medline: 15208226]
39. Costanzo C, Verghese A. The physical examination as ritual: social sciences and embodiment in the context of the physical examination. Med Clin North Am 2018 May;102(3):425-431. [doi: 10.1016/j.mcna.2017.12.004] [Medline: 29650064]
40. Polinski JM, Barker T, Gagliano N, Sussman A, Brennan TA, Shrank WH. Patients' satisfaction with and preference for telehealth visits. J Gen Intern Med 2016 Mar;31(3):269-275 [FREE Full text] [doi: 10.1007/s11606-015-3489-x] [Medline: 26269131]
41. Alepis E, Lambrinidis C. M-health: supporting automated diagnosis and electonic health records. Springerplus 2013 Dec;2(1):103 [FREE Full text] [doi: 10.1186/2193-1801-2-103] [Medline: 23556144]
42. Chowdhury M, Birns J, Rudd A, Bhalla A. Telemedicine versus face-to-face evaluation in the delivery of thrombolysis for acute ischaemic stroke: a single centre experience. Postgrad Med J 2012 Mar;88(1037):134-137. [doi: 10.1136/postgradmedj-2011-130060] [Medline: 22267526]
43. Hill E, Whitehead M, MacInnes B, Ellis G, Talbot A, Brodie F, et al. The first 100 thrombolysis cases in a novel Scottish mesh telestroke system. Scott Med J 2013 Nov;58(4):213-216. [doi: 10.1177/0036933013507868] [Medline: 24215039]
44. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of ehealth interventions: systematic review of the literature. J Med Internet Res 2018 May 01;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
45. Hanlon P, Daines L, Campbell C, McKinstry B, Weller D, Pinnock H. Telehealth interventions to support self-management of long-term conditions: a systematic metareview of diabetes, heart failure, asthma, chronic obstructive pulmonary disease, and cancer. J Med Internet Res 2017 May 17;19(5):e172 [FREE Full text] [doi: 10.2196/jmir.6688] [Medline: 28526671]
46. Morey S, Hamoodi A, Jones D, Young T, Thompson C, Dhuny J, et al. Increased diagnosis and treatment of hepatitis C in prison by universal offer of testing and use of telemedicine. J Viral Hepat 2019 Jan;26(1):101-108. [doi: 10.1111/jvh.13017] [Medline: 30315691]
47. Morris J, Campbell-Richards D, Wherton J, Sudra R, Vijayaraghavan S, Greenhalgh T, et al. Webcam consultations for diabetes: findings from four years of experience in Newham. Pract Diab 2017 Mar 20;34(2):45-50. [doi: 10.1002/pdi.2078]
48. Dorsey ER, Topol EJ. State of telehealth. N Engl J Med 2016 Jul 14;375(2):154-161. [doi: 10.1056/NEJMra1601705] [Medline: 27410924]
49. Norman S. The use of telemedicine in psychiatry. J Psychiatr Ment Health Nurs 2006 Dec;13(6):771-777. [doi: 10.1111/j.1365-2850.2006.01033.x] [Medline: 17087682]
50. Hill A. Telerehabilitation in Scotland: current initiatives and recommendations for future development. Int J Telerehabil 2010;2(1):7-14 [FREE Full text] [doi: 10.5195/ijt.2010.6055] [Medline: 25945168]
51. Banerjee M, Chakraborty S, Pal R. Diabetes self-management amid COVID-19 pandemic. Diabetes Metab Syndr 2020;14(4):351-354 [FREE Full text] [doi: 10.1016/j.dsx.2020.04.013] [Medline: 32311652]
52. Pappas Y, Seale C. The opening phase of telemedicine consultations: an analysis of interaction. Soc Sci Med 2009 Apr;68(7):1229-1237. [doi: 10.1016/j.socscimed.2009.01.011] [Medline: 19201514]
53. Zhao Y, Li K, Zhang L. A meta-analysis of online health adoption and the moderating effect of economic development level. Int J Med Inform 2019 Jul;127:68-79. [doi: 10.1016/j.ijmedinf.2019.04.015] [Medline: 31128834]
54. Giunti G, Goossens R, De Bont A, Visser JJ, Mulder M, Schuit SC. The need for sustainable teleconsultation systems in the aftermath of the first COVID-19 wave. J Med Internet Res 2020 Oct 05;22(10):e21211 [FREE Full text] [doi: 10.2196/21211] [Medline: 32997642]
55. James HM, Papoutsi C, Wherton J, Greenhalgh T, Shaw SE. Spread, scale-up, and sustainability of video consulting in health care: systematic review and synthesis guided by the NASSS framework. J Med Internet Res 2021 Jan 26;23(1):e23775 [FREE Full text] [doi: 10.2196/23775] [Medline: 33434141]

Abbreviations

CCA: canonical correlation analysis
DM: diabetes mellitus
DNA: did not attend
ICT: information and communication technology
IoU: intention of use
TAM: technology acceptance model
TH: teleconsultations at home
UTAUT: unified theory of acceptance and use of technology

Edited by A Kushniruk; submitted 10.02.21; peer-reviewed by R Ciorap, R Suomi, P Murphie; comments to author 10.07.21; revised version received 22.07.21; accepted 23.07.21; published 23.11.21.

Please cite as:
Rego N, Pereira HS, Crispim J
Perceptions of Patients and Physicians on Teleconsultation at Home for Diabetes Mellitus: Survey Study
JMIR Hum Factors 2021;8(4):e27873
URL: https://humanfactors.jmir.org/2021/4/e27873
doi:10.2196/27873
PMID:34817394

Original Paper

User Interactions With Health Insurance Decision Aids: User Study With Retrospective Think-Aloud Interviews

Wayne C W Giang[1], BASc, MASc, PhD; Emma Bland[1], BS; Jeffrey Chen[1]; Coralys M Colón-Morales[1], BSc, MS; Michelle M Alvarado[1], BSc, ME, PhD

Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, United States

Corresponding Author:
Wayne C W Giang, BASc, MASc, PhD
Department of Industrial and Systems Engineering
University of Florida
303 Weil Hall
P.O. Box 116595
Gainesville, FL, 32611-6595
United States
Phone: 1 352 294 7729
Email: wayne.giang@ise.ufl.edu

Abstract

Background: Two barriers to effective enrollment decisions are low health insurance literacy and lack of knowledge about how to choose a plan. To remedy these issues, digital decision aids have been used to increase the knowledge of plan options and to guide the decision process. Previous research has shown that the way information is presented in a decision aid can impact consumer choice, and existing health insurance decision aids vary in their design, content, and layout. Commercial virtual benefits counselors (VBCs) are digital decision aids that provide decision support by mimicking the guidance provided by an in-person human resources (HR) counselor, whereas more traditional HR websites provide information that requires self-directed navigation through the system. However, few studies have compared how decision processes are impacted by these different methods of providing information.

Objective: This study aims to examine how individuals interact with two different types of health insurance decision aids (*guided* VBCs that mimic conversations with a real HR counselor and *self-directed* HR websites that provide a broad range of detailed information) to make employer-provided health insurance decisions.

Methods: In total, 16 employees from a local state university completed a user study in which they made mock employer-provided health insurance decisions using 1 of 2 systems (VBC vs HR website). Participants took part in a retrospective think-aloud interview, cued using eye-tracking data to understand decision aid interactions. In addition, pre- and postexperiment measures of literacy and knowledge and decision conflict and usability of the system were also examined.

Results: Both the VBC and HR website had positive benefits for health insurance knowledge and literacy. Previous health insurance knowledge also impacted how individuals used decision aids. Individuals who scored lower on the pre-experiment knowledge test focused on different decision factors and were more conflicted about their final enrollment decisions than those with higher knowledge test scores. Although both decision aids resulted in similar changes in the Health Insurance Literacy Measure and knowledge test scores, perceived usability differed. Website navigation was not intuitive, and it took longer to locate information, although users appreciated that it had more details; the VBC website was easier to use but had limited information. Lower knowledge participants, in particular, found the website to be less useful and harder to use than those with higher health insurance knowledge. Finally, out-of-pocket cost estimation tools can lead to confusion when they do not highlight the factors that contribute to the cost estimate.

Conclusions: This study showed that health insurance decision aids help individuals improve their confidence in selecting and using health insurance plans. However, previous health insurance knowledge plays a significant role in how users interact with and benefit from decision aids, even when information is presented in different formats.

(JMIR Hum Factors 2021;8(4):e27628) doi:10.2196/27628

KEYWORDS

insurance, health; health benefits plans; employee; decision aids; cognitive ergonomics; human factors engineering

Introduction

Background

Health insurance enrollment is a complex decision based on many factors, such as price, product attributes, and current health status, and can significantly impact a person's health and financial circumstances [1]. Making an informed decision depends on a person's knowledge, literacy, cognitive skills, and confidence to carry out said tasks [2]. Despite the consequences of this choice, only 4% of adults in the United States understand basic health insurance terminology and often get overwhelmed by the complexity of the decision [3,4]. In addition, health insurance plans can be challenging to understand, especially if the decision-maker has limited financial or health insurance literacy [2,5]. Previous research has shown that the way information is presented in the decision aid can impact consumer choice [6]. Factors such as the order in which plans are presented, word choice and symbol use, and difficulty in finding information can significantly affect trust in the information and consumer choice [1,7]. Therefore, understanding how individuals interact with sources of health insurance information is a key component in improving informed decision-making.

Over 55% of Americans receive health insurance from their employer [8]. Some employers have used virtual benefits counselors (VBCs) to provide further decision support for their employees. VBCs are designed to provide guided support by mimicking a one-on-one interaction with a human resources (HR) representative. Although there has been limited research on health insurance decision aids [6,9,10], VBCs are a relatively new product for supporting health insurance choices that combine access to tools, such as cost estimators, and further guidance and recommendations presented through a conversational interface. The effects of this more guided approach to decision-making are still not well understood, and little research has examined how guided systems affect consumer health insurance decision-making when compared with traditional self-directed methods such as websites.

VBCs may be of particular benefit to low-literacy consumers, as previous research has shown that these consumers often confuse health insurance concepts [2]. Kodagoda et al [11] found that users with low reading literacy, numeracy, and digital literacy tend to end information searches early (due to perceived completion of task), take longer to complete the tasks, and have less directed searching strategies than high-literacy users. In addition, low-literacy users are less able to predict where information would be on a website accurately, are less able to find information on websites, and are less likely to verify the information found. VBCs guide users through the enrollment process and provide relevant information and recommendations on the basis of cost calculations using a conversational interface. Consequently, they have the potential to help users, especially those with lower literacy, make informed health insurance enrollment decisions. The guided decision support provided by a VBC system has the potential to improve a user's ability to find relevant information and ensure that they consider important factors while making their enrollment decisions.

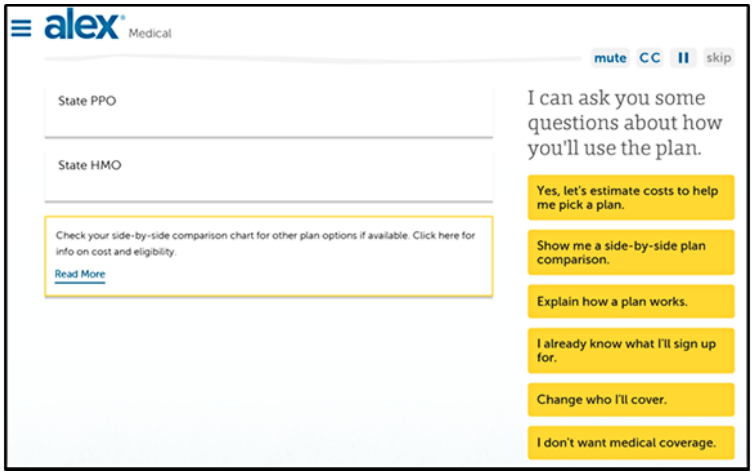
User interactions with health insurance information and digital decision aids, such as VBCs and HR websites, are likely to be impacted by their incoming knowledge and previous health insurance use [3]. Furthermore, if participants are able to become more knowledgeable and literate about health insurance information, they are likely to become more informed and confident decision-makers. These changes may have influenced the factors considered during the decision process. However, few studies have directly compared VBCs and HR websites, particularly for employer-provided plans. Most studies have also largely relied on reviews of health insurance enrollment data sets [12] or web-based evaluations of decision aids [10] to evaluate the effectiveness of these decision aids, which makes it more difficult to understand the user's decision process as they interact with these tools. Thus, this exploratory study uses a think-aloud method to understand how an individual's interaction with the guided VBC decision aid versus self-directed information provided on HR websites influences the user's decision process and measures that may impact the final decision quality: health insurance knowledge, literacy, decision conflict, system usability, and decision processes.

Health Insurance Decision Aids

Virtual Benefits Counselors

This study uses Alex, a VBC created by Jellyvision Lab Inc, which was customized to the specific plans provided by the employer. Alex uses a conversational question-and-answer interface with colorful animations, text, and a fully voiced personality (Figure 1). The conversation guides the interactions of the user and helps to structure the decision process. Alex also interjects at different points to provide definitions or clarifications of the information provided.

Figure 1. Screenshot of the Jellyvision Lab Inc's Alex interface.



HR Website

HR benefits websites provide a self-directed experience in which users navigate freely between different pages. During the study, the state university's HR benefits website had health information distributed across two areas: a general benefits section and a dedicated section on health insurance. Information about eligibility, comparison charts between plans, and enrollment processes can be found on these pages. The HR website also provides links to the state's health insurance website, where details about the different plan options (Health Maintenance Organization and Preferred Provider Organization) including costs (ie, deductibles, premiums, copays, and coinsurance), network size, and coverage were presented using digital brochures and tables. Overall, the website provided detailed information that was distributed nonlinearly across multiple pages and lacked the cost estimation tools found in the VBC.

Methods

Participants

Participants were recruited as part of a survey studying sources of health insurance information used to make enrollment decisions at a local employer (a university campus) that has been reported elsewhere [13,14]. Links to the survey were distributed to the staff and faculty, resulting in a total of 140 complete responses. Of these 140 responses, 113 (80.7%) indicated an interest in participating and were contacted for recruitment. A total of 16 employees enrolled and completed the user study and were randomly assigned to either the VBC or HR website. All participants indicated that they had primary (11/16, 69%) or shared responsibility (5/16, 31%) for health care decisions in their household. Data collection was impacted by COVID-19 during the data collection phase, which resulted in a smaller participant sample than that initially planned. This study was approved by the local institutional review board.

Experimental Task

To understand how guided and self-directed information support affects informed health insurance decision-making, a *mock health insurance enrollment task* along with a *retrospective think-aloud method* was used. Retrospective think-alouds can detect issues during user interactions and help encourage participants to verbalize comments about their thoughts and interactions with the system [15].

Mock Health Insurance Enrollment Task

Participants were asked to make a mock health insurance enrollment decision for the upcoming year for their household. They were provided with a decision aid to assist them with this task: either the VBC or HR website. Participants were asked to stay within the bounds of the provided system and were directed to return to the system if they exceeded these bounds (eg, left the HR website to use Google).

Think-Aloud Interview

A retrospective think-aloud method was used to understand the participants' motivations and strategies when navigating through their assigned system. Participants were asked to explain their thought process, the information they were looking for, and anything they were confused or unsure about while watching a video of their gaze behavior while using the system, which was captured using an eye tracker. The experimenters occasionally prompted participants to verbalize their thoughts and reasoning behind decisions throughout the interview and ask for clarifications when required. The interviews were recorded and transcribed using a Health Insurance Portability and Accountability Act–compliant transcription service.

Procedure

The experiment consisted of five phases. (1) *Training*: participants were introduced to the purpose of the study, eye-tracking equipment, and think-aloud methodology. They then went through a training session where they were familiarized with the eye tracking and think-aloud process with a simple decision task. (2) *Pre-experiment questionnaire*: participants were given a pre-experiment questionnaire that measured their health insurance literacy and health insurance knowledge; (3) *mock health insurance enrollment task,* as described in the experimental tasks. (4) *Postexperiment questionnaire*: after making their enrollment decision, participants were again asked about their health insurance literacy and knowledge. Participants were also asked to fill a decision conflict scale and to rate the usefulness and ease of use of the system; (5) *a retrospective think-aloud interview,* as described in the experimental tasks. The experiment was

conducted by a trained graduate and undergraduate research assistant in an office-like environment. The experiment took approximately 90 minutes.

Experimental Design and Measures

The main independent variable was the decision aid system used to assist with the mock enrollment decision, either the VBC or HR website. The response measures were the thematic analysis of the think-aloud interviews and questionnaire data delivered before and after engaging with the system.

The thematic analysis of the think-aloud interviews allowed for the identification of *decision factors* discussed by participants during their use of the 2 systems. These factors provide insight into the variables considered by the participants while making enrollment decisions. Think-aloud interviews were also used to identify themes about *how participants interacted* with the guided VBC and self-directed HR website decision aid systems.

Four sets of questionnaire data were also examined to help understand the participants' health insurance literacy, confidence in decision-making, and perceived usability of the 2 systems.

The *Health Insurance Literacy Measure (HILM)* is a 21-item self-report questionnaire that asks participants to assess their self-efficacy in four subcomponents of health insurance related to confidence and likelihood of demonstrating health insurance literate behaviors: *confidence in choosing a plan*, *comparing plans*, *confidence in using a plan*, and *being proactive* when using a plan. Participants rated each item on a 7-point scale (from "1-extremely low/extremely unlikely" to "7-extremely high/extremely likely"), which was averaged to calculate a score for each category.

The *knowledge tests* included seven true or false questions about different health insurance concepts and definitions. One test was adapted from a previous study on health insurance decision aids by Politi et al [10]. A second version of the test was created with a similar difficulty. The order of the tests was counterbalanced.

The *SURE* (Sure of myself; Understand information; Risk-benefit ratio; Encouragement) *measure* is a series of 4 yes or no questions designed to measure decisional conflict, with higher scores indicating less decisional conflict [16].

Participants were also asked to rate the *usefulness* and *ease of use* of the system on a 10-point scale (from "1-not useful at all/not easy to use at all" to "10-extremely useful/extremely easy to use").

Apparatus

Participants used a 15-inch laptop with an attached mouse to navigate through the decision aids. Eye-tracking data were collected using a Tobii Pro Nano screen-based eye tracker, and a retrospective think-aloud was facilitated using Tobii Pro Lab software [17].

Data Analysis

Emergent Themes Analysis (ETA) was used to identify the factors that each participant mentioned during their enrollment decision while using the 2 decision aid systems. ETA has previously been used to understand decision processes [18,19] and user interactions with decision aids [11,20]. The process started with identifying broad themes or conceptually related topics found within the transcripts through an initial high-level reading of the data. Three researchers (WCWG, JC, and MMA) completed this process and identified a number of common themes that were mentioned by many of the 16 participants. These themes were consolidated through a card-sort. This analysis was supplemented with observations and quotations about the users' strategies to engage with the decision aids. Participants were also divided into 2 groups—those who came into the experiment with lower health insurance knowledge (scores<6/7) and those with higher health insurance knowledge, and this variable was used in subsequent analyses.

Owing to the small sample size, exploratory data analysis was conducted on the questionnaire data to better understand the effects of interacting with the decision aids on health insurance literacy and knowledge, SURE scores, and usefulness and ease of use ratings. These descriptive quantitative data were further supported using excerpts from think-aloud interviews.

Data from 2 participants were partially impacted by data recording issues. The pre-experiment questionnaire data for 1 VBC participant was lost, and their data were excluded from the analysis of the HILM, knowledge test, SURE scores, and usefulness and ease of use ratings. An HR website participant had eye-tracking data recording issues during the mock health insurance enrollment task, resulting in a think-aloud interview based on a video of the interactions rather than prompted by eye-tracking data; these data were kept within the data set.

Results

Respondent Demographics and Characteristics

The demographics of the 16 participants are presented in Table 1. Participants were predominantly female (12/16, 75%), and the majority were married or in a domestic partnership (10/16, 63%). However, most of the participants came from small households of either 1 or 2 individuals (11/16, 69%). Finally, most participants chose the Health Maintenance Organization plan (10/16, 63%) over the Preferred Provider Organization plan. Both plans had similar desirability as they covered similar procedures and services and had the same monthly premium but differed in terms of network, deductible, and coinsurance or copays.

Table 1. Participant demographics and plan choice.

Demographics	VBC[a] participants (n=8), n (%)	Human resources website participants (n=8), n (%)	Total (n=16), n (%)
Gender			
Female	7 (88)	5 (63)	12 (75)
Male	1 (12)	3 (37)	4 (25)
Age (years)			
18-24	1 (12)	1 (12)	2 (12)
25-34	4 (50)	3 (37)	7 (44)
35-44	2 (25)	2 (25)	4 (25)
45-54	1 (12)	1 (12)	2 (12)
55-66	0 (0)	1 (12)	1 (6)
Marital status			
Married	4 (50)	6 (75)	10 (63)
Not married	4 (50)	2 (25)	6 (37)
Average time since hire (years)			
<1	2 (25)	2 (25)	4 (25)
2-5	4 (50)	3 (38)	7 (44)
>5	2 (25)	2 (25)	4 (25)
Number of additional family members covered in the plan			
0	3 (38)	2 (25)	5 (31)
1	3 (38)	3 (38)	6 (38)
2-4	2 (25)	3 (38)	5 (31)
Selected plan			
HMO[b]	6 (75)	4 (50)	10 (63)
PPO[c]	2 (25)	4 (50)	6 (37)

[a]VBC: virtual benefits counselor.

[b]HMO: Health Maintenance Organization.

[c]PPO: Preferred Provider Organization.

Health Insurance Literacy and Knowledge

Table 2 shows the pre- and postexperiment scores for the 4 dimensions of the HILM and the knowledge test for the VBC and the HR website. Across the sample, participants tended to rate their confidence and likelihood of performing health insurance literate behaviors as higher than neutral, with the scales relating to likelihood of performing health literate behaviors (eg, comparing plans or being proactive) scoring higher than the confidence scales (eg, confidence in choosing or confidence in using). Figure 2 shows the differences in the HILM scores for each participant after interacting with the decision aid. Across all four subcomponents, the differences appeared to be similar for the VBC and the HR website. However, most participants reported an increase in their confidence in choosing (8/15, 53%) and using (10/15, 67%) their plans, which agreed with the larger magnitude in scores seen in Table 2 (Choosing: Δ=0.43; Using: Δ=0.44). In contrast, the 2 HILMs related to health insurance literate behaviors had more variable results after interaction with both the VBC and the HR website.

Surprisingly, the trend of the data suggested that the *comparing plans* literacy subcomponent decreased after interacting with the decision aids, and the magnitude of this decrease was larger for the VBC. The interview data provided additional evidence for this, with multiple VBC participants commenting that they were often confused by the outputs of the cost estimation tool as the Health Maintenance Organization and Preferred Provider Organization plans resulted in very similar out-of-pocket costs. Thus, participants felt that the VBC recommendations were not useful because they did not understand why a plan was selected. For instance, 1 VBC participant commented:

> *And this is when I came up to the not very helpful conclusion that the plans are essentially the same. [...] I was a little surprised. I figured there would be a little bit more difference between the two programs. And then, I said, "It's not really helpful," because it's not, I mean, they're so similar and they don't do a good job in this [VBC] system, in my opinion, of explaining what the differences are and really highlighting those differences, because there are*

> *differences, and I think they really should emphasize those instead of basically saying, "Oh, well, they're the same.*" [[ID03, VBC]]

This quote highlights that participants may not have been using the information provided by the VBC's cost estimation tools to their full extent. As mentioned previously, the cost estimation tool allows users to estimate their yearly out-of-pocket costs by guiding users by estimating health insurance use throughout the year and automatically calculating the final costs for the different plan options. One method for using such a tool is to allow users to estimate the effects of different health insurance use scenarios. For example, participants may want to see how the costs of each plan change if they require major surgery in the upcoming year. Participants appeared to be aware of this possibility, with one participant commenting during their interactions with the cost estimation tool, "But I think, yes of course you want to include like at least one [specialist visit], because at least maybe show you the cost differential." [ID12, VBC]

By exploring these different outcomes, users may be better able to understand the differences between the provided plans and how uncertainty in their own health insurance use might impact the overall costs. However, none of the participants in our VBC condition used the cost estimation tool. Instead, some participants commented that they would have liked to see how their use choices influenced the VBC's cost estimation:

> *That's exactly what I would have liked, is a breakdown of all those things that I chose, like, which one of them [was contributing to the costs] because then I could be like, "Well, I put two ER costs, but that's not really the biggest deal here." I think that would have been helpful.* [ID08, VBC]

These results suggest that cost estimation tools should highlight how changes in health insurance use would impact the final out-of-pocket costs, and this design change may lead to users considering a more diverse set of health care use scenarios.

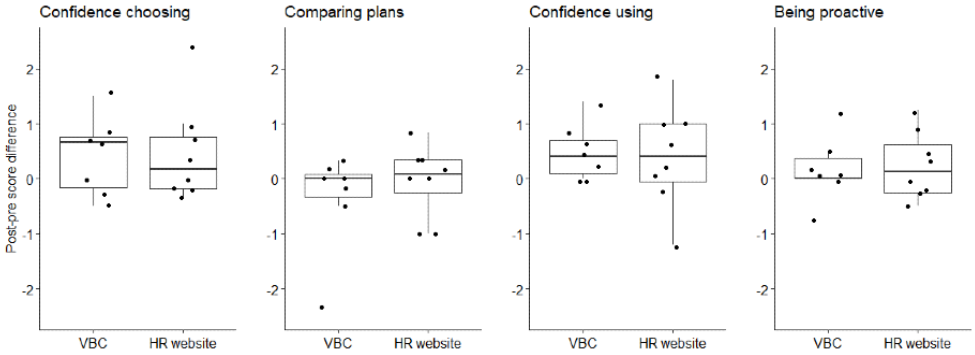
As shown in Table 2, participants in the sample had relatively high knowledge test scores (out of 7) before interacting with the decision aids (VBC: median 6, range 3-7; HR website: median 5.5, range 4-7). Their knowledge test scores increased after interacting with the decision aids (VBC: median 6, range 4-7; HR website: median 7, range 6-7). The average improvement in knowledge test scores was similar between the VBC (Δ=1.14) and HR website (Δ=1.13). These results suggest that both decision aid systems improve health insurance knowledge.

Table 2. Sample means and SDs for pre- and postexperiment Health Insurance Literacy Measure and knowledge test scores.

Dimensions and test	VBC[a] (n=7), mean (SD)		Human resources website (n=8), mean (SD)	
	Pre-experiment	Postexperiment	Pre-experiment	Postexperiment
Confidence in choosing	4.10 (1.14)	4.5 (1.03)	4.56 (1.37)	5.02 (1.36)
Comparing plans	5.60 (1.14)	5.24 (1.62)	5.52 (1.40)	5.48 (1.27)
Confidence in using	3.97 (1.23)	4.46 (0.97)	4.45 (1.65)	4.85 (1.61)
Being proactive	5.54 (0.87)	5.71 (0.81)	6.09 (0.76)	6.34 (0.33)
Knowledge test (%)	71.4 (23.2)	87.7 (15.3)	80.3 (15.1)	96.4 (6.6)

[a]VBC: virtual benefits counselor.

Figure 2. Differences in post- and pre-experiment Health Insurance Literacy Measure scores. HR: human resources; VBC: virtual benefits counselor.



Decision Factors

The thematic analysis in Table 3 identifies several decision factors considered by the participants during their enrollment process. The types of factors considered by participants in both the VBC and HR website conditions were relatively similar, with network size, use costs, and coverage costs being the factors considered most frequently, and the ease of use and understanding plans being the factors considered least frequently.

Table 3. Decision factors considered by participants during the enrollment process compared across decision aid (virtual benefits counselor [VBC] vs human resources [HR] website) and pre-experiment knowledge test (KT) scores (low vs high).

Decision factor	Definition	Example quotes	VBC (n=8), n (%)	HR website (n=8), n (%)	Low KT score (n=7), n (%)	High KT score (n=8), n (%)
Network size	Consideration of the size of the network or whether their providers or specialists were within the network	"HMO is not a good thing. I mean, it's not good. I knew it, but based on your information, again, it's not good for out of network at all. But still there's a high probability that I will be out of network, so I want to have some coverage for that specific thing." [ID01, VBC]	6 (75)	7 (88)	6 (86)	6 (75)
Premium costs	The monthly premium costs for each plan	"Those [premiums] are really the costs that I care about more, because if I'm going to be paying it [...] every single month I'll want to know that." [ID06, VBC]	5 (63)	6 (75)	5 (71)	5 (63)
Use costs	Copays, coinsurance, and deductible costs	"It's talking about the lower deductible at $250 a person or $500 a family." [ID07, HR website]	6 (75)	7 (88)	4 (57)	8 (100)
Coverage costs	Costs associated with specific treatments, procedures, or care and whether they are covered by the plan	"Generally, I have drug costs [...] so, now I went into the plan because I wanted to see if there was anything on the cost of specialty drugs. [...] Everything I was reading about confirmed that it was the right type of drug [...] it just didn't tell me the cost. I'm doing the same thing here looking for the cost of the specialty drug, which is 100% covered under the HMO." [ID15, HR website]	7 (88)	6 (75)	5 (71)	7 (88)
Estimated out-of-pocket costs	The total yearly estimated out-of-pocket costs	"What I was trying to do overall was figure out my total cost for the year in each plan. So then, I had to figure out $500 a year and that's when I did the premium $420 versus $50 like what the $35 a month savings was. The high-deductible that was again like - the savings would be not enough because the out-of-pocket max is so high." [ID15, HR website]	6 (75)	6 (75)	6 (86)	5 (63)
Ease of use of plan	How easy it was to use the plan to access and pay for care	"I'm just more familiar with the HMO and all of our – we haven't had an issue where we couldn't really find a provider that was under a plan because, you know, should you have chances, it's pretty – yeah, it's pretty accessible." [ID16, VBC]	3 (38)	3 (38)	2 (29)	4 (50)
Ease of understanding the plan	How easy it was to understand the plan	"It seems to me that the HMO is, you know, easier to understand. I tend, you know, not to trust… these two [Plan Name] plans, because it is confusing." [ID02, HR website]	2 (25)	2 (25)	3 (43)	1 (13)

These results were surprising because of the very different information presentation methods and tools provided in the guided VBC and self-directed HR website. The VBC explicitly offers a series of factors to consider through cost estimation tools and highlights important plan features in comparison tables. In contrast, the HR website requires individuals to identify and seek relevant information and tools. The fact that participants using both systems were able to identify similar types of decision factors suggests that participants in this study had some idea of what elements they should look at while searching for health insurance and agreed with the HILM and knowledge test scores discussed previously. They did not necessarily depend on the guidance provided by the VBC to identify new, personally significant decision factors.

However, the ETA helped identify 2 elements that might make it difficult for participants to use the decision factors, even though they had identified them. First, participants may struggle to obtain accurate information about each factor; either they were not able to quickly locate the information in the system (a common complaint for the HR website condition because of the layout of the pages or not knowing the correct keyword to search for, eg, "My routine bloodwork, I have a hard time kind of finding bloodwork, but eventually, I figure out what term to search for. I just used a search function to figure out both of the keywords that I had in my mind" [ID05, HR website]) or the information was not provided at a level of detail that the participant desired (a common issue in the VBC condition, see the coverage cost quote in Table 3 for a contrasting HR website example). Second, participants may not know how to use the identified factors to make a final decision. The VBC benefited from a cost estimation tool that helped users estimate their yearly costs based on user-provided estimates of their health insurance

use for the coming year and calculate a final value. However, participants in the HR website condition were required to perform these calculations. Although many of the participants in the HR website condition attempted to do so (eg, estimated out-of-pocket cost quote in Table 3), these calculations are likely to be more difficult and more error-prone without decision support.

Table 3 also shows which factors participants with higher and lower prior knowledge discussed. Many participants with higher pre-experiment knowledge test scores mentioned use costs (eg, copays, coinsurance, and deductibles) during their think-aloud interviews and considered ease of plan use during their decision-making process. In contrast, participants with lower knowledge test scores discussed how easy it was to understand a health care plan during the think-aloud interviews. These results further emphasize the vital role of health insurance literacy in how individuals interact with decision aids, regardless of the guided or self-directed presentation methods.

Decision Conflict

The SURE score represents the amount of decision conflict that a decision-maker has about their final choice. Any score below 4 indicated a conflict or lack of comfort with the final decision. As seen in Table 4, participants in both the VBC and HR website conditions had some decision conflict, although those in the HR website condition appeared to have less conflict. This was surprising because the VBC provided explicit guidance for the decision process, whereas the HR website did not. However, 1 participant stated that it was this additional guidance that made them less confident in their final decision when speaking about the VBC's cost estimation tool:

> *Because it makes me have to stop and think about these estimates of how much care I think I would need [...] versus if I was just looking at the flat numbers in my head, I probably wouldn't say, "Oh, how many of these visits do you do?" [...] I just look at the flat numbers and say, "Well, this one is cheaper." And I go with that.* [[ID12, VBC]]

Table 4. SURE (Sure of myself, Understand information, Risk-benefit ratio, Encouragement) scores for system type (virtual benefits counselor [VBC] vs human resources [HR] website) and lower versus higher preinteraction knowledge test (KT) scores.

SURE score	VBC (n=7), n (%)	HR website (n=8), n (%)	Lower KT scores (n=7), n (%)	Higher KT scores (n=8), n (%)
0	2 (29)	0 (0)	2 (29)	0 (0)
1	0 (0)	2 (25)	2 (29)	0 (0)
2	1 (14)	1 (13)	0 (0)	2 (25)
3	1 (14)	0 (0)	1 (14)	0 (0)
4	3 (43)	5 (63)	2 (29)	6 (75)

In contrast, participants in the HR website condition may have considered fewer unknown situations or novel factors when making their decision. Many of the decision factors considered by the HR website participants were those that had already been previously considered before interacting with the decision aid or an ongoing medical condition, for example:

> *Like as my [spouse] and I get older, [spouse] is needing some surgeries. My children are in sports. We travel quite a bit for sports. You just never know when someone is going to get hurt.* [ID07, HR website]

As these factors were produced by the participants on their own volition, they may have resulted in less conflict when making the decision.

The SURE scores of participants who had lower pre-existing knowledge (eg, lower scores on the pre-experiment knowledge test) also appeared to have more decision conflict than those who had higher scores (Table 4). Similar to participants in the HR website condition, participants who had higher pre-existing knowledge also tended to have specific factors that they searched for during the decision process, regardless of the system (eg, "I did choose to look at the out-of-network versus the in-network because it's kinda topical right now for me. So, my [spouse] is gonna go to New Mexico for a two – like, two plus months [...] and [they] currently got some health problems." [ID16, VBC]).

Participants with higher knowledge also tended to be more active users of their current plans and were much more familiar with the health insurance enrollment decision. Thus, regardless of the system type, previous knowledge plays a significant role in confidence in their final decision.

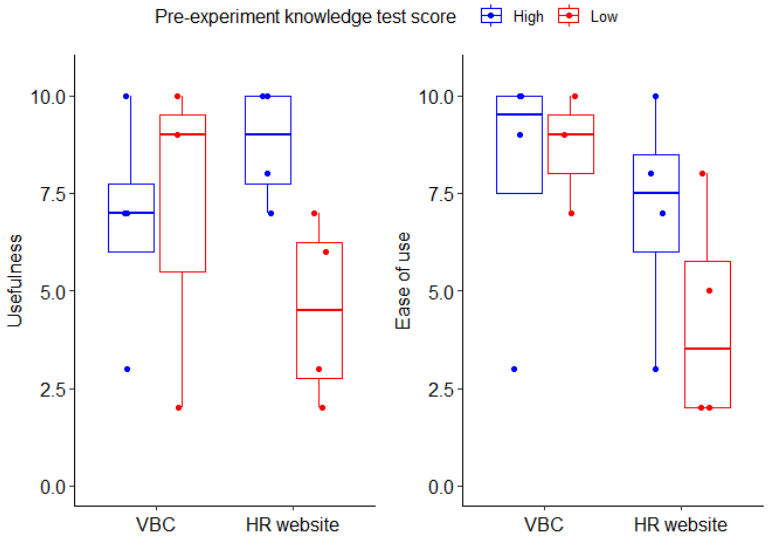
Usefulness and Ease of Use

Finally, participants were asked to rate the usefulness and ease of using the two systems. Figure 3 shows the ratings across the two systems for participants who scored low and high on the pre-experiment knowledge test. The sample data suggest that pre-existing knowledge about health insurance played a role in how the participants rated the usability of the VBC and HR website. A number of participants stated that they believed that the VBC would be useful for those who had lower health insurance literacy (eg, "I mean I'm familiar a little bit with insurance like I definitely am not a person who knows nothing. So, I can see at how it can be helpful for somebody with no knowledge of health care." [ID16, VBC]), whereas those with higher pre-existing knowledge commented that the VBC did not have all the required details (eg, the VBC was "not tied in. I still have to go to these crazy insurance websites, actually get to the information, where I really want to be at that detail I want." [ID12, VBC]). This trend was supported by the usefulness ratings with lower knowledge participants rating slightly higher usefulness than those with higher knowledge. A larger difference existed for the HR website condition, where the higher knowledge participants likely benefited from the

extra information available on the website. However, those with lower knowledge had more difficulties in finding information and making decisions. For example, 1 participant commented that they felt that information on the website was repetitive and that it was hard to understand where they needed to go:

> *I would say it's a bit cumbersome, and it creates confusion. But, you know, if you search carefully, you are able to find some useful information, but it takes time and effort. I would say, you know, this website assistant can be greatly simplified, and there's a lot of information redundancy [...] I really need someone, you know, who really knows it, who can provide me with some advice.* [ID02, HR website]

Figure 3. Participant ratings of usefulness and ease of use across system and pre-experiment knowledge test scores. HR: human resources; VBC: virtual benefits counselor.



For ease of use, both higher and lower knowledge participants rated the VBC system as easy to use. The VBC system ratings were also higher than those of the HR website. Individuals with lower health insurance knowledge rated the HR website as more difficult to use than those with higher pre-existing knowledge. This was likely due to differences in how the 2 groups navigated and searched for information. Participants with lower health insurance knowledge tended to describe their search process as reading through the different options laid out on each page before deciding on which link to click on next (eg, "I started with the [State health insurance webpage], just because it's on top, reading to find clickable items, hit my benefits, and scrolling down to just kind of see what this page had." [ID11, HR website]). However, even higher knowledge participants noted issues with the organization of the information and knowing where to go next (eg, "To me, it's just not obvious. I am not sure I need to be clicking multiple times to get where I need to go. I feel like on the very first page I should be able to click health insurance plans." [ID07, HR website]).

Overall, these findings agree with the other measures; both the guided VBC system and the self-directed HR website were seen as useful and had positive benefits to users. However, the guided VBC system experience was perceived to be easier to use than the self-directed HR website, particularly for participants who had lower pre-experiment health insurance knowledge.

Discussion

Principal Findings

This study examined how individuals use 2 different types of health insurance decision aids to make enrollment decisions: (1) a guided VBC system that walked users through factors that are important in health insurance decision-making and provide support for terminology and definitions and (2) a more traditional digital source of health insurance information, an HR website that provides educational information and brochures but requires self-directed navigation through the system. We contrasted these 2 decision aid systems on measures that may impact decision quality: health insurance knowledge, literacy, decision factors, decision conflict, and usability.

The results showed that both types of health insurance decision aids had positive benefits for health insurance knowledge and literacy. Previous health insurance knowledge played an important role in how individuals used the 2 health information decision aids. Individuals with lower pre-experiment knowledge test scores focused on different decision factors and were more conflicted about their final enrollment decisions than those with higher knowledge test scores. Furthermore, although both decision aids resulted in similar changes in the HILM and knowledge test scores, differences exist for the usefulness and ease of use of the 2 systems. HR website navigation was not intuitive, and it took longer to locate information, although users appreciated that it had more details; the VBC system was easier to use but had limited details with some users, indicating that the HR website was still needed as a supplementary companion. Lower knowledge participants, in particular, found the HR website to be less useful and harder to use than those with higher health insurance knowledge. Finally, decision aid tools, such as out-of-pocket cost estimation tools, can lead to confusion when they do not highlight which factors of each plan contribute to the cost estimate. Users wanted a more robust tool that

showed the cost breakdown and could help them explore how their use estimates influence cost estimation.

Comparison With Previous Work

One surprising finding in this study was that the VBC and the HR website had similar effects on health insurance knowledge and literacy. This contrasts with the results of Politi et al [10], who compared a custom-built decision aid (consisting of education, cost estimations, and recommendations) with a traditional government website. Their decision aid resulted in higher literacy and knowledge than websites. This difference could be due to the smaller sample size (n=16 vs n=328) or participant demographics. In addition, our study participants had employer-provided health insurance, whereas the majority in the study by Politi et al [10] were uninsured. A study by Vardell [21] on new employees choosing health insurance found that first-time decision-makers tended to have lower HILM than other participants. Owing to our participants' previous experience with the plans, they may have had less ability to benefit from a health insurance decision aid. Furthermore, participants in both conditions used similar decision factors, even though the VBC system provided more guidance about the factors to consider. Future work will be required to expand our exploratory study to a more diverse population.

However, interacting with either system appeared to have positive benefits for measures that may lead to more informed decision-making. Participants scored better on the knowledge test and felt more confident about choosing and using their plans (2 subcomponents of health literacy) after using the decision aids. These are positive outcomes given the low health insurance literacy and health insurance plan selection issues previously found in the literature [2,5,21-23]. However, 1 trend in both the literacy measure and the think-aloud interviews was confused about comparing plans. Our participants had difficulty understanding what made the plans unique. This suggests that current decision aids fail to help users develop a mental model of how each plan works. Mental models are a type of internal representation that helps simulate different future outcomes [24,25]. More fully developed mental models may help users explore possible future scenarios and better understand each plan's strengths and weaknesses. Interestingly, none of our participants used the cost estimation tool to explore these possibilities. Decision aids may require additional guidance for users to explore *edge-cases* rather than just focusing on the most likely scenarios. This type of support is likely more challenging to implement in a self-directed system such as a webpage than a guided VBC decision aid and should be explored in future work.

Finally, the results suggest that pre-existing knowledge may have one of the most significant impacts on the decision-making process regardless of the type of decision aid provided. In our study, participants with higher pre-experiment knowledge test scores had fewer decision conflicts. They also focused on essential decision factors, such as the plan's ease of use, in contrast to those with lower knowledge test scores, who focused on how easy it was to understand a plan. Our findings agree with previous research indicating that individuals with low HILM scores or limited experience with health insurance decision-making (eg, new employees or young adults) will struggle with the decision process, how to interact with the system, and are more likely to make mistakes when choosing coverage [21,22,26]. Our results also suggest that the unguided nature of the HR website made it more difficult for those with low health insurance knowledge to use and benefit from because they do not have a strategy for searching for relevant information and combining this information together to make a final decision. The VBC, on the other hand, had similar benefits for both low and high knowledge participants, with the main drawback being the lack of detailed information. Thus, VBCs may fulfill their intended purpose of helping those with low health insurance literacy but should be used with a variety of different sources of health insurance information.

Limitations

A few limitations exist that may impact the generalizability of the results of this study. The study had a small sample of only 16 participants, as recruitment for in-person human subject experiments was impacted by COVID-19. Although these participants generated a large set of think-aloud interview data (approximately 4 hours of audio recordings), the small sample size for the knowledge test and literacy measures made it difficult to compare the VBC and HR websites using inferential statistical analysis. Instead, our analysis focused on interview data and an exploratory analysis of the trends in the questionnaire data. Furthermore, the employees in our sample were full-time employees at a local state university. These included both staff and faculty and may not be fully representative of employees at other types of employers in terms of demographics, health insurance knowledge, and experience with technology. Future work should examine how differences between employee characteristics at different employers may impact user interactions with health insurance decision aids.

Conclusions

In conclusion, this study showed that health insurance decision aids help individuals improve their knowledge about health insurance and their confidence in selecting and using health insurance plans. In addition, previous health insurance knowledge played a significant role in how users interacted with and benefited from decision aids. Although the study participants indicated that both the VBC and HR website appeared to have a similar effect on these HILMs and decision factors considered, participants perceived the VBC system as easier to use. In contrast, participants with lower prior knowledge appeared to struggle with using the HR website, resulting in lower perceived usefulness and ease of use.

This study's thematic analysis identified important decision factors among the study participants. Once again, the specific decision aid did not strongly impact the relative importance of the decision factors. However, participants with low health insurance knowledge felt more conflicted about their final mock decisions. In addition, they discussed the health plan's ease of use and use costs less frequently than others, but they also placed more value on how well the digital aids helped them understand the plan. Finally, more research is required on (1) how decision aids affect mental models of health insurance plans, (2) how decision aids affect user decision strategies and

information-seeking strategies, and (3) the development of more robust cost estimation tools that help users differentiate plans for *edge-cases* and out-of-pocket costs.

Conflicts of Interest

None declared.

References

1. Taylor E, Carman K, Lopez A, Muchow A, Roshan P, Eibner C. Consumer Decisionmaking in the Health Care. Santa Monica, California: RAND Publications; 2016.
2. Politi MC, Kaphingst KA, Kreuter M, Shacham E, Lovell MC, McBride T. Knowledge of health insurance terminology and details among the uninsured. Med Care Res Rev 2014 Feb 24;71(1):85-98. [doi: 10.1177/1077558713505327] [Medline: 24163306]
3. Lubalin JS, Harris-Kojetin LD. What do consumers want and need to know in making health care choices? Med Care Res Rev 1999;56(Suppl 1):67-102. [doi: 10.1177/107755879905600 1S04] [Medline: 10354679]
4. 4 health insurance terms 96% of Americans don't know Internet. Policygenius. 2018. URL: https://www.policygenius.com/health-insurance/learn/health-insurance-literacy-survey/ [accessed 2021-09-20]
5. Bhargava S, Loewenstein G. Choosing a health insurance plan: complexity and consequences. JAMA 2015 Dec 15;314(23):2505-2506. [doi: 10.1001/jama.2015.15176] [Medline: 26670967]
6. Wong CA, Kulhari S, McGeoch EJ, Jones AT, Weiner J, Polsky D, et al. Shopping on the public and private health insurance marketplaces: consumer decision aids and plan presentation. J Gen Intern Med 2018 Aug;33(8):1400-1410 [FREE Full text] [doi: 10.1007/s11606-018-4483-x] [Medline: 29845467]
7. Ye Y. Correlates of consumer trust in online health information: findings from the health information national trends survey. J Health Commun 2011 Jan;16(1):34-49. [doi: 10.1080/10810730.2010.529491] [Medline: 21086209]
8. Keisler-Starkley K, Bunch L. Health Insurance Coverage in the United StatesInternet. United States Census Bureau. URL: http://www.census.gov/content/dam/Census/library/publications/2015/demo/p60-253.pdf [accessed 2021-09-20]
9. Barnes AJ, Hanoch Y, Rice T, Long SK. Moving beyond blind men and elephants: providing total estimated annual costs improves health insurance decision making. Med Care Res Rev 2017 Oct;74(5):625-635. [doi: 10.1177/1077558716669210] [Medline: 27624636]
10. Politi MC, Kuzemchak MD, Liu J, Barker AR, Peters E, Ubel PA, et al. Show me my health plans: using a decision aid to improve decisions in the federal health insurance marketplace. MDM Policy Pract 2016 Jul;1:11 [FREE Full text] [doi: 10.1177/2381468316679998] [Medline: 28804780]
11. Kodagoda N, Wong W, Kahan N. Identifying information seeking behaviours of low and high literacy users: combined cognitive task analysis. In: Proceedings of the 9th International Conference on Naturalistic Decision Making. 2009 Jun Presented at: 9th International Conference on Naturalistic Decision Making; June, 2009; BCS London URL: http://eprints.mdx.ac.uk/4751/ [doi: 10.14236/ewic/ndm2009.58]
12. Handel BR, Kolstad JT. Health insurance for "humans": information frictions, plan choice, and consumer welfare. Am Econ Rev 2015 Aug 01;105(8):2449-2500. [doi: 10.1257/aer.20131126] [Medline: 29546969]
13. Colón-Morales C, Giang W, Alvarado M. Informed decision-making for health insurance enrollment: survey study. JMIR Form Rese 2021 Aug 12;5(8):e27477. [doi: 10.2196/preprints.27477] [Medline: 34387555]
14. Colón Morales C, Giang WCW, Alvarado M. Effectiveness of health insurance enrollment decision aids. In: Proceedings of the 2020 IISE Annual Conference 2020.. 2020 Presented at: IISE Annual Conference; May 30-June 2,2020; Florida, United States p. 1276.
15. Elbabour F, Lecturer A, Alhadreti O, Mayhew P, Lecturer S. Eye tracking in retrospective think-aloud usability testing: is there added value? J Usability Stud 2017 May;12(3):110. [doi: 10.14236/ewic/hci2016.101]
16. Légaré F, Kearing S, Clay K, Gagnon S, D'Amours D, Rousseau M, et al. Are you SURE?: assessing patient decisional conflict with a 4-item screening test. Can Fam Physician 2010 Aug;56(8):e308-e314 [FREE Full text] [Medline: 20705870]
17. Tobii PA. Tobii Pro Lab. Tobii Pro. URL: https://www.tobiipro.com/product-listing/tobii-pro-lab/ [accessed 2021-09-06]
18. Harenčárová H. Structured analysis of critical decision method data – emergency medicine case study. J Hum Affairs 2015 Oct 06;25(4):459. [doi: 10.1515/humaff-2015-0036]
19. Wong BL, Blandford A. Analysing ambulance dispatcher decision making: trialing emergent themes analysis. In: Proceedings of the HF2002 Human Factors Conference: Design for the whole person - integrating physical, cognitive and social aspects : a joint conference of the Ergonomics Society of Australia (ESA) and the Computer Human Interaction Special Interest Group (CHISIG). 2002 Presented at: HF2002 Human Factors Conference: Design for the whole person; Nov 25-27, 2002; Melbourne ,Australia.
20. Kodagoda N, Wong BL. Effects of Low & High Literacy on User Performance in Information Search and Retrieval. In: Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction,

BCS HCI .–181. 2008 Presented at: 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction, BCS HCI .–181; Sep 1-5, 2008; Liverpool, United Kingdom. [doi: 10.14236/ewic/hci2008.17]
21. Vardell E. Health insurance literacy: how people understand and make health insurance purchase decisions. ProQuest Dissertations and Theses. 2017. URL: https://core.ac.uk/download/pdf/210596286.pdf [accessed 2021-09-06]
22. Wong CA, Asch DA, Vinoya CM, Ford CA, Baker T, Town R, et al. Seeing health insurance and HealthCare.gov through the eyes of young adults. J Adolesc Health 2015 Aug;57(2):137-143. [doi: 10.1016/j.jadohealth.2015.04.017] [Medline: 26092178]
23. Bartholomae S, Russell MB, Braun B, McCoy T. Building health insurance literacy: evidence from the Smart Choice Health Insurance™ program. J Fam Econ Iss 2016 Feb 20;37(2):140-155. [doi: 10.1007/s10834-016-9482-7]
24. Wilson JR, Rutherford A. Mental models: theory and application in human factors. Hum Factors 1989 Dec 01;31(6):617-634. [doi: 10.1177/001872088903100601]
25. Thatcher A, Greyling M. Mental models of the internet. Int J Ind Ergon 1998 Nov;22(4-5):299-305. [doi: 10.1016/s0169-8141(97)00081-4]
26. Barcellos SH, Wuppermann AC, Carman KG, Bauhoff S, McFadden DL, Kapteyn A, et al. Preparedness of Americans for the Affordable Care Act. Proc Natl Acad Sci U S A 2014 Apr 15;111(15):5497-5502 [FREE Full text] [doi: 10.1073/pnas.1320488111] [Medline: 24706843]

Abbreviations

ETA: Emergent Themes Analysis
HILM: Health Insurance Literacy Measure
HR: human resources
SURE: Sure of myself, Understand information, Risk-benefit ratio, Encouragement
VBC: virtual benefits counselor

Edited by A Kushniruk; submitted 31.01.21; peer-reviewed by E Vardell, M Politi; comments to author 22.02.21; revised version received 02.07.21; accepted 04.07.21; published 26.10.21.

Please cite as:
Giang WCW, Bland E, Chen J, Colón-Morales CM, Alvarado MM
User Interactions With Health Insurance Decision Aids: User Study With Retrospective Think-Aloud Interviews
JMIR Hum Factors 2021;8(4):e27628
URL: https://humanfactors.jmir.org/2021/4/e27628
doi:10.2196/27628
PMID:34698647

Original Paper

User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service

Hedvig Aminoff[1], MSc; Sebastiaan Meijer[1], PhD; Kristina Groth[2], PhD; Urban Arnelo[3,4], MD, PhD

[1]Department of Biomedical Engineering and Health Systems, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm, Sweden

[2]The Center for Innovation, Karolinska University Hospital, Stockholm, Sweden

[3]CLINTEC Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden

[4]Department of Surgical and Perioperative Sciences, Umeå University, Umeå, Sweden

Corresponding Author:
Hedvig Aminoff, MSc
Department of Biomedical Engineering and Health Systems
School of Engineering Sciences in Chemistry, Biotechnology and Health
KTH Royal Institute of Technology
Hälsovägen 11 C
Huddinge
Stockholm, 14157
Sweden
Phone: 46 8 790 60 00
Email: hedvigam@kth.se

Related Article:

Comment in: https://humanfactors.jmir.org/2022/2/e36681

Abstract

Background: Teleguidance, a promising telemedicine service for intraoperative surgical consultation, was planned to scale up at a major academic hospital in partnership with 5 other hospitals. If the service was adopted and used over time, it was expected to provide educational benefits and improve clinical outcomes during endoscopic retrograde cholangiopancreatography (ERCP), which is a technically advanced procedure for biliary and pancreatic disease. However, it is known that seemingly successful innovations can play out differently in new settings, which might cause variability in clinical outcomes. In addition, few telemedicine services survive long enough to deliver system-level outcomes, the causes of which are not well understood.

Objective: We were interested in factors related to usability and user experience of the telemedicine service, which might affect adoption. Therefore, we investigated perceptions and responses to the use and anticipated use of a system. Technology acceptance, a construct referring to how users perceive a technology's usefulness, is commonly considered to indicate whether a new technology will actually be used in a real-life setting. Satisfaction measures were used to investigate whether user expectations and needs have been met through the use of technology. In this study, we asked surgeons to rate the perceived usefulness of teleguidance, and their satisfaction with the telemedicine service in direct conjunction with real-time use during clinical procedures.

Methods: We designed domain-specific measures for perceived usefulness and satisfaction, based on performance and outcome measures for the clinical procedure. Surgeons were asked to rate their user experience with the telemedicine service in direct conjunction with real-time use during clinical procedures.

Results: In total, 142 remote intraoperative consultations were conducted during ERCP procedures at 5 hospitals. The demand for teleguidance was more pronounced in cases with higher complexity. Operating surgeons rated teleguidance to have contributed to performance and outcomes to a moderate or large extent in 111 of 140 (79.3%) cases. Specific examples were that teleguidance was rated as having contributed to intervention success and avoiding a repeated ERCP in 23 cases, avoiding 3 PTC, and 11 referrals, and in 11 cases, combinations of these outcomes. Preprocedure beliefs about the usefulness of teleguidance were

generally lower than postprocedure satisfaction ratings. The usefulness of teleguidance was mainly experienced through practical advice from the consulting specialist (119/140, 85%) and support with assessment and decision-making (122/140, 87%).

Conclusions: Users' satisfaction with teleguidance surpassed their initial expectations, mainly through contribution to nontechnical aspects of performance, and through help with general assessment. Teleguidance shows the potential to improve performance and outcomes during ERCP. However, it takes hands-on experience for practitioners to understand how the new telemedicine service contributes to performance and outcomes.

(JMIR Hum Factors 2021;8(4):e30867) doi:10.2196/30867

KEYWORDS

telemedicine; user experience; satisfaction; technology acceptance; usability; perioperative; surgery; consultation; surgeons; performance; evaluation; teleguidance; telehealth; telemedicine implementation

Introduction

Overview

A telemedicine service for intraoperative surgical consultation during advanced gastrointestinal endoscopy was seen as having the potential to increase the quality and safety of procedures, and there was hope that the technique could be used in other areas of medicine. The service, called teleguidance, had shown success in a feasibility study [1], and health economic modeling showed its potential for improving clinical and economic outcomes [2]. Teleguidance practice was to be scaled up to 4 additional hospitals (5 in total including the first hospital), and efforts were made to understand the context into which the telemedicine intervention was introduced [3], and potential users' attitudes toward the service prior to implementation [4]. This paper describes an investigation of practitioners' experience of real-time use of teleguidance, based on surgeons' expectations of how the service might contribute to performance and outcomes in a particular procedure, and their satisfaction with teleguidance immediately after the procedure.

Background

Teleguidance is a professional-to-professional telemedicine service for video collaboration during a highly specialized endoscopic procedure called endoscopic retrograde cholangiopancreatography (ERCP).

ERCP is a technically advanced procedure for examination, sampling, and interventions in the complicated ductal structures of the gall bladder, pancreas, and liver; for example, to remove or alleviate blockages caused by tumors in the liver. Today, ERCP is a common procedure, and when successful, it can quickly relieve very painful and serious conditions; it is sometimes a prerequisite for consecutive procedures. However, ERCP is a complex, high-risk, collaborative task and a highly technical specialty with a long learning curve. Learning the perceptual and motor skills to control the equipment and interpreting the guiding video and x-ray images requires considerable practice and skill, combined with clinical decision-making based on careful weighing of risks and benefits.

Clinical practitioners had expressed positive expectations that teleguidance could contribute to the quality and safety of procedures but also concerns that teleguidance might disrupt work practices [4]. The telemedicine service was designed through participatory, user-centered development [5], and there were hopes that by providing remote intraoperative consultation, teleguidance could contribute to learning and improved performance, which could enable practitioners at smaller hospitals to provide more highly specialized procedures. A feasibility study had reported clinical benefits [1], and health economic modeling showed the potential for positive clinical and economic outcomes [2]. A decision was made to scale up teleguidance to 4 additional hospitals, with an intention to generate additional evidence for the benefits of the practice.

The Need for Expertise in ERCP

Increasing therapeutic use of ERCP and increasing procedural complexity has raised the level of expertise required for ERCP [6]. At smaller hospitals in Sweden, many individual ERCP practitioners and clinics have an annual procedural volume which is below the recommendations for sustaining and advancing skill [7].

Traditionally, advanced surgical skills are learned by working together with experienced surgeons, progressing from shadowing to increasingly independent work with hands-on training and mentorship. Once proficiency is gained, a certain procedural volume is generally considered necessary to sustain newly acquired skills, develop experience, and keep up with new technical advances, and high procedural volume is associated with fewer adverse events [8].

Studies have shown that there are large variations in the quality of ERCP procedures at different clinics [9], and failure to cannulate the desired duct or post-ERCP pancreatitis is common but has serious consequences [10]. Repeated unsuccessful attempts to cannulate the correct duct play a significant role for complications [11], and cannulation failure can lead to a decision to abort a procedure, causing a subsequent delay in treatment, or conversion to more invasive procedure, such as percutaneous transhepatic cholangiography (PTC) [12]. Post-ERCP pancreatitis and other serious adverse events such as bleeding or perforation, have been found to be more common when the procedure was performed by less experienced practitioners [13].

It has been suggested that ERCP specialists with lower levels of expertise should not attempt complex or difficult ERCP cases without the assistance of a more experienced endoscopist [12], and that serious outcomes can be avoided if there is an option to cooperate with other highly specialized colleagues in the case of adverse events [6].

At larger hospitals, a practitioner needing advice during a difficult procedure often has a colleague ERCPist on call to advise or assist during a difficult procedure. This is not always the case at smaller hospitals, and practitioners have commonly used the telephone when they needed advice. Remote surgical guidance in ERCP, *teleguidance*, via videoconferencing and simultaneous transfer of high-quality surgical imaging was developed to enhance this practice to help develop and sustain the expertise of individual practitioners at lower-volume centers [1].

Remote Surgical Consultation and Mentoring

Services similar to teleguidance, such as surgical telementoring and remote surgical guidance, have previously been developed to support education and address knowledge gaps [14-16] and increase access to highly specialized treatment in remote or low-case-load facilities [17-20]. They have, for example, been used in trauma and emergency medicine [21] and in laparoscopic and open surgery [22].

However, for the potential benefits of any telemedicine service to be fulfilled over time, a service also needs to become a part of regular practice [23]. Despite a wide range of apparent benefits, there is limited evidence for the educational benefits of telementoring [15], and this way of working has generally failed to become a daily tool in clinical workflows [15,16,18,24]. While a wide range of barriers to telemedicine implementation and adoption have been identified [25], how these individual factors contribute to assimilation and sustainable use of telemedicine in real-life practice is not well understood [26]. However, feedback about real-time use of telemedicine systems can be a way to gain a better image of the factors that can affect clinical outcomes and adoption [27-29].

In evaluations of the impact of telemedicine services on clinical outcomes, telemedicine can be defined as a complex intervention [30]; as such, there is value in complementing traditional clinical assessments of effectiveness with qualitative studies of user perceptions and experience [29,31]. According to ISO 9241-11, usability is the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [32]. Usability evaluation can include combinations of objective measures of effectiveness (eg, successful task completion rate) and efficiency (eg, task completion time), and subjective measures of satisfaction [33], which can be a valuable a part of an effort to understand the likelihood of acceptance and use of a telemedicine system.

Perceived Usefulness and Satisfaction

Telemedicine research has shown that users' perceptions of a technology's usefulness are a main significant predictor of acceptance [26]. The concept of technology acceptance is widely used in telemedicine research [26] but originated in information systems research [34], where behavioral theory and methodologies have generated a large body of research about how users' attitudes toward a technology influence subsequent adoption and use [35]. The technology acceptance model (TAM) [36] conceptualizes acceptance as an evaluative process, where technology use can be predicted or explained on the basis of psychometric measures of users' expectations about perceived usefulness; that is, how using a technology will impact job effectiveness, efficiency, and performance [37].

Satisfaction is also a considered a key component for telemedicine success and is often included in the evaluation of telemedicine services [38]. Telemedicine satisfaction studies have often reported favorable results [39,40]; nonetheless, methodological issues often make it difficult to interpret or compare findings [41], and it is often unclear what satisfaction measures actually demonstrate [42]. In health services research, satisfaction measures may implicitly refer to patient satisfaction with treatment or care [43,44], which contrasts with measures of satisfaction with the use of a technology; that is, whether user expectations and needs have been met through the use of technology [32,35,45,46].

This indicates a need for careful definition and operationalization of measures of perceived usefulness and satisfaction when collecting feedback about the real-time use of telemedicine, to ensure a match between the constructs and the technology involved, the users, and the context in which the telemedicine service is being introduced.

Aims and Objectives

Telemedicine acceptance and adoption are not well understood [47,48], and many telemedicine services fail to be adopted, despite their apparent value; this lack of successful implementation is so common it has been described a "paradox of telehealth" [49-51]. However, information systems research has shown that users' expectations of a technology's usefulness is an important determinant for adoption, and that satisfaction measures can be used to investigate whether user expectations of using a technology have been met.

When teleguidance was scaled up, we were interested in how clinical practitioners experienced the usefulness of the telemedicine service, and if the service lived up to these expectations, as this could provide insight into the factors shaping the success of the intervention. We were provided with an opportunity to gather data in direct conjunction with teleguidance sessions, and wanted to investigate whether clinical practitioners expected teleguidance to contribute to a specific case, and whether these expectations were met during the teleguidance session. These subjective measures of anticipated usefulness and satisfaction with using teleguidance are expected to provide knowledge about central user perceptions and user experience, which can influence the implementation, adoption, and use of remote surgical consultation.

Methods

Methods Overview

Participating surgeons at the central hospital and the 5 participating remote sites filled in case report forms (CRF) in conjunction with each teleguided ERCP procedure. These data were passed in raw form to us. In the following sections, the design of the CRFs and the rationale for data collection, which underlies how we operationalized the constructs perceived usefulness and satisfaction, is described.

Design of the CRF

Two paper-based CRFs were designed to gather pre- and postprocedure data. The guiding surgeon at the central site registered patient- and case-related data, mainly to ensure correct orientation about the case to be guided and to provide ratings of the technical quality of image transmission. The remote surgeons were asked to register patient- and intervention-related data and a rating of the level of complexity of each case. In addition, they were asked to provide subjective ratings of their estimated need and expectations for consultation during the procedure, and to report technical issues and their experience of how teleguidance contributed to performance and outcomes.

The Guiding Surgeons' CRF

Prior to a teleguidance session, the guiding specialist requires basic information about the case and the patient, and this was communicated either by telephone or through the videoconferencing system. The guider was asked to register the data necessary to ensure a correct understanding of the needs and potential risks during consultation: if it is an emergency or planned elective procedure, the patient's gender, age, and whether ERCP has been performed previously. In addition, knowledge about whether the remote procedure was to be conducted with sedation or general anesthesia provided information about the patient's orientation, which has consequences regarding the interpretation of the transmitted image. The indications and aims of the procedure provided fundamental information to the guiding specialist about what was to be done during the ERCP.

The Remote Surgeons' CRF

The CRF contained data entry fields about clinical indications, the aims and success of the procedure, and 30-day follow-up items about complications, consecutive procedures, and health economic data. The CRF also included 2 subjective preprocedure rating items about the benefits the participants were hoping for and the problems they were hoping to avoid through teleguidance. After teleguided procedures, satisfaction with teleguidance was measured through participants' ratings of the ways in which teleguidance contributed to their performance and to the outcomes of the procedure.

Appropriate indication, cannulation rate, stone extraction success rate, stent insertion success rate, and post-ERCP pancreatitis frequency are evidence-based, prioritized quality indicators for ERCP [12]. Therefore, we considered these factors relevant for perceived usefulness and investigated whether participants believed that teleguidance could contribute to cannulation, stone extraction, and stent insertion. We also investigated whether practitioners considered teleguidance to support clinical assessment and decision-making, and help avoid additional interventions or referrals.

Performance and outcome measures such as cannulation frequency or complication frequency are not straightforward to interpret: more complex cases generally have a higher risk for complications [13]. Therefore, we also wanted to obtain information about the clinical difficulty of the procedure. This was measured through a preprocedure case complexity rating, with 4 predefined categories of clinical contexts, techniques, and anatomical or pathological features (Multimedia Appendix 1).

Participants were also asked to grade cannulation difficulty after each teleguided procedure in accordance with the 5-5-2 principle defined by the European Society of Gastrointestinal Endoscopy [52]. The number of earlier ERCPs the patient has undergone and the characteristics of the papilla are also associated with cannulation outcomes [53] and were therefore also included in the CRF.

Multimedia Appendix 2 shows the CRF items related to perceived usefulness, prior to teleguidance sessions, and satisfaction immediately experience after teleguided sessions.

Procedure

Each participating hospital received a utility cart equipped with the necessary components to transmit endoscopic and fluoroscopic images. The cart was also equipped with a camera and microphone to capture images and sounds from the operation theater.

The teleguidance equipment had one video and one content channel, which meant that the participants had to choose from among endoscopy, fluoroscopy, or a view of the operating room. The remote party controlled switching between imaging, and would change upon request from the consulting surgeon. Audio communication was possible throughout the session.

The remote sites used the following teleconference systems: a mobile Polycom Practitioner Cart HDX unit (Polycom) equipped with a 26-inch LCD screen, a high-definition video camera, stereo speakers, and microphones (Figure 1). At the University hospital, a Polycom HDX 4500 desktop videoconference system was used to provide guidance from either an office near to the endoscopy suite in which the ERCP interventions were usually performed, or an office equipped with multiple videoconference systems set up specifically for teleguidance.

All communication passed through Sjunet, a secure, IP-based broadband network for Swedish health care providers (Sjunet).

Figure 1. The remote clinics used a Polycom Realpresence Practitioner cart 8000 (A). At the central site, the consulting surgeon used a Polycom HDX4500 desktop video conferencing system with a touch screen control (B).

There was no function for telestration (annotating live video content telestration) during scaling up of teleguidance. During the design of the teleguidance solution, a prototype for telestration (Multimedia Appendix 3) had been tested in a few scenarios. It was hoped that a function for graphical annotation would improve shared understanding in tasks such as localization of the point of entry to the common bile duct. However, the results indicated that users on both ends of the teleguidance session were distracted by the design and function of this particular telestration solution. Our findings echoed those of a more well-designed study [54], and a decision was made not to further develop this function.

Written instructions, contact details to the guiding practitioners, and technical support, and a protocol for establishing a connection among the hospitals were also provided (Multimedia Appendix 4). Clinical and technical staff received a tutorial.

The operating surgeons and the consulting surgeon used the telephone to agree on the timing for the teleguidance session days ahead, or in some cases, immediately before a procedure. The remote sites initiated the teleguidance sessions.

CRFs were distributed to each site to be filled in on paper by the operating endoscopist and submitted to a coordinating research nurse at the central site.

Participants were instructed to teleguide as many ERCP cases as possible during the study period and not to select cases. Operating surgeons were asked to book teleguidance sessions in advance by telephone, but there was also the option to call in direct conjunction with a procedure.

Two senior ERCP experts at a high-volume tertiary referral ERCP clinic provided remote consultation via teleguidance.

Sample

ERCP procedures at 5 district hospitals received teleguidance from a tertiary referral center. In total, 142 teleguided procedures are included in the sample. The average duration of ERCP procedures was 53 minutes (range 10-224 minutes, median 45 minutes). The average duration of teleguidance sessions was 43 minutes (range 4-186 minutes, median 35 minutes).

The most common indications were biliary and pancreatic stones (56/142, 39%) and icterus (33/142, 23%); the most common aim for procedures were ERC stone extraction (71/142, 50%) and ERC offload (59/142, 42 %) (Multimedia Appendix 5). A total of 75 of 142 (53%) patients were female, and 67 of 142 (47%) were male. The age range was between 18 and 91 years, the mean age of female patients was 67 years and that of men was 67 years. Furthermore, 43 of 142 (30%) cases were emergency interventions, while 93 of 142 (66%) were elective (5/142, 4% were not classified). Additional details about the patients and case complexity ratings are shown in Multimedia Appendix 6.

In total, 14 ERCPists participated at the remote sites. Table 1 shows the level of experience among participants at the remote sites. All 5 novices with low experience (>200 ERCP) progressed to an expert level of 500-1000 ERCP procedures during this period. The distribution of cases across hospitals and practitioners is shown in Multimedia Appendix 7.

Table 1. Level of experience among participants at remote sites (n=11).

Participants	Guided sessions (n=142), n (%)
Novice (<200 ERCP[a])	14 (9.9)
Novice (200-500 ERCP)	33 (23.2)
Expert (500-1000 ERCP)	87 (61.3)
Expert (>1000 ERCP)	8 (5.6)

[a]ERCP: endoscopic retrograde cholangiopancreatography.

The remote sites reported technical issues in 26 of 142 (11%) cases; however, these were problems experienced by the consulting surgeon at the central site. This did, however, cause inconvenience at the remote sites, since the remote sites had to conduct troubleshooting in these cases.

The central site reported technical issues in 34 of 142 (24%) cases. In total, 16 of 34 (47%) of the reported problems regarded acoustic feedback between microphones on the teleguidance cart and microphones in the operation theater at one of the hospitals. In 9 of 34 (26%) cases, there were problems with pixelated image quality or problems with hue. In 5 of 34 (15%) cases, the consulting surgeon could see the endoscopic video, but there were intermittent problems with transfer on fluoroscopy.

In some cases, this was resolved by restarting the connection; in some cases, medical technicians at the remote sites provided assistance and resolved problems; for example, by changing video graphics array cables between monitors. While this caused some delays, the teleguidance sessions proceeded despite the technical issues. There were 4 cases of postoperative complications (4/142, 2.8%).

Results

Perceived Usefulness of Teleguidance

Perceived usefulness was measured on the basis of surgeons' ratings of their expectations of how teleguidance might contribute to procedures. Table 2 and Figure 2 show how surgeons rated their anticipated demand for teleguidance prior to specific procedures: in 58 of 139 (41%) procedures, surgeons expected to have use for teleguidance; 42 of 139 (30%) reported that they did not; and 38 of 139 (27%) were unsure.

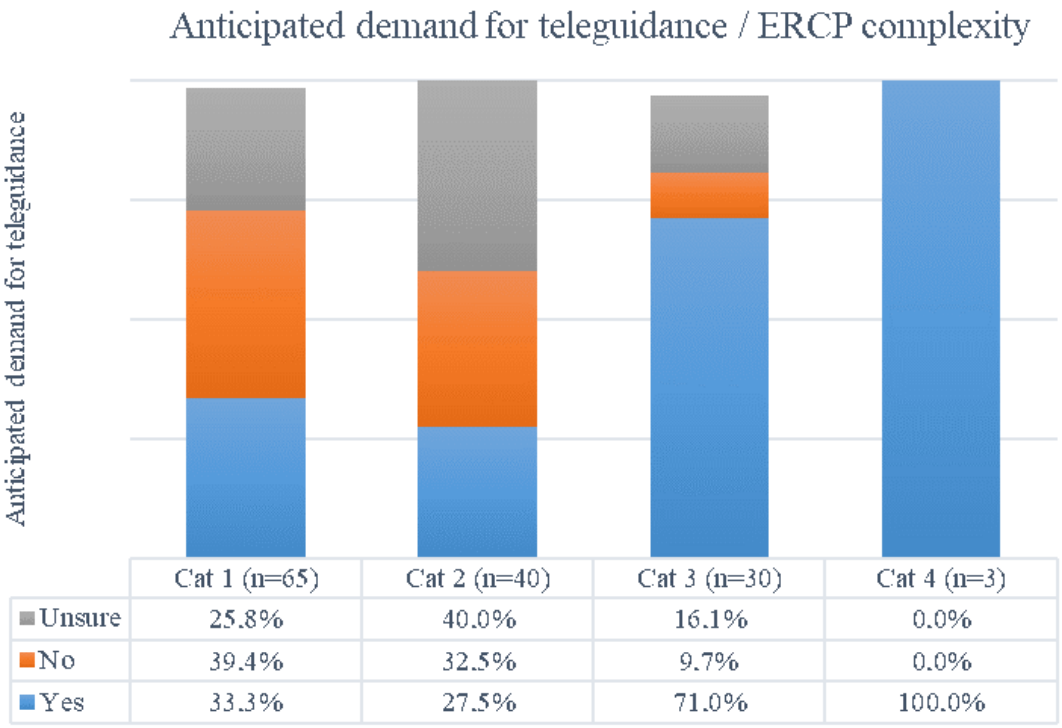
In 20 cases, the option to use teleguidance had affected the decision to perform the procedure in question.

The operating surgeon's level of expertise did not significantly affect the anticipated demand for teleguidance (r=0.008, P>.001) (Multimedia Appendix 8).

Table 2. Anticipated demand for teleguidance.

Response	Surgeons, n (%)
Yes	58 (42.0)
No	42 (30.4)
Unsure	38 (27.5)

Figure 2. Anticipated demand for teleguidance/endoscopic retrograde cholangiopancreatography complexity. ERCP: endoscopic retrograde cholangiopancreatography.



However, the demand for teleguidance was more pronounced in cases with higher complexity according to the 4-category rating scale (Multimedia Appendix 1). Case complexity and the anticipated demand for teleguidance (Table 2) showed a significant linear relationship (r=–0.229, P<.001).

Participants expressed a higher anticipated demand for teleguidance for certain intervention goals; however, we did not observe any significant relationships. In ERC stone extractions, only 21/69 (30%) rated a need for teleguidance while there was more demand in the other procedures, especially regarding pancreatic procedures: ERC offload (28/58, 47.5%), endoscopic retrograde cholangiography (ERC) biopsy (12/21, 57%), and the pancreatic procedures ERP stone extraction (6/11, 54.5%) and ERP offload (8/11, 72.7%) (Multimedia Appendix 9).

Similarly, participants expressed a higher anticipated demand for teleguidance for certain indications (Multimedia Appendix 10). The demand for teleguidance was high (>40%) for acute pancreatitis (7/15, 46.7 %), chronic pancreatitis (10/12, 83.3%), primary sclerosing cholangitis (PSC) (5/5, 100%), and strictures with unknown causes (5/9, 55.6%).

In 82 of 142 (57.7%) cases, surgeons reported expectations to avoid certain situations through teleguidance (Multimedia Appendix 11). The most frequent situation they hoped to avoid was having to repeat the ERCP, which is commonly owing to a failure to cannulate (42/82, 51.2%).

In 67 of 142 (53%) cases, the surgeons expressed an expectation to receive support with specific tasks during the procedure (Multimedia Appendix 12). In total, 10 of 67 (14.9%) hoped to receive support with cannulation; 4 of 67 (6%) hoped to receive support with the placement of a stent; 6 of 67 (9%) hoped to receive support with the removal of a stone; 27 of 67 (34%) hoped to receive support with clinical assessment; and 21 of 72 (28%) hoped to receive support with combinations of these tasks.

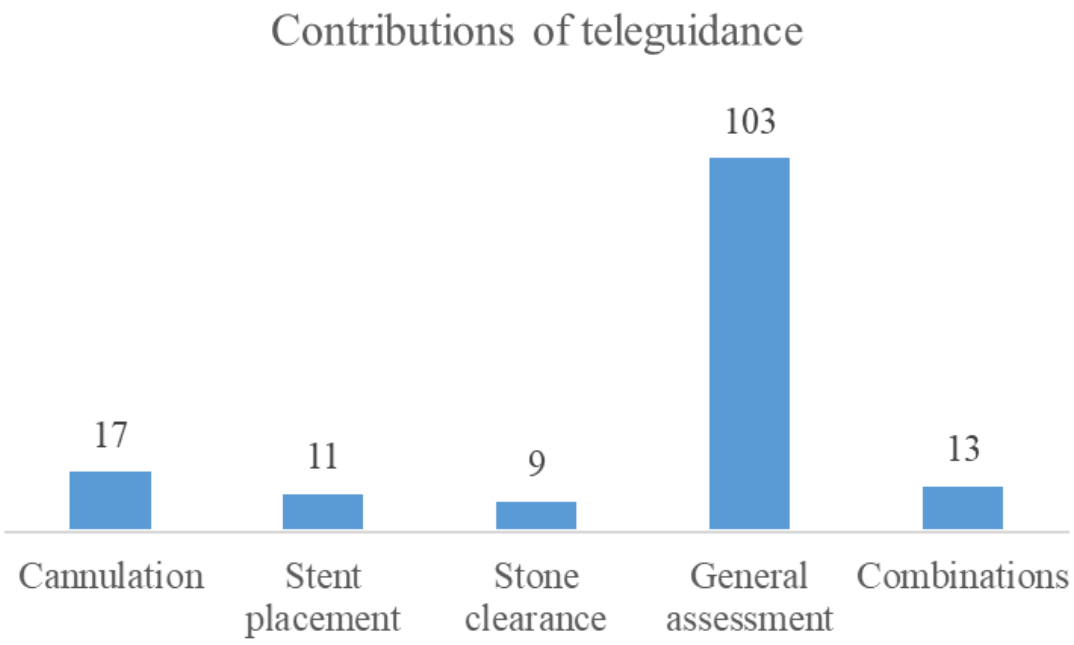
Satisfaction With Teleguidance

Satisfaction was measured after procedures, through ratings of how teleguidance contributed to performance and outcomes.

The operating endoscopists rated teleguidance to have contributed value to a moderate or large extent (rating value 3 and 4) in 111 of 140 (79.3%) cases (Multimedia Appendix 13). In 16 of 26 (61%) cases where the operating surgeon reported technical problems, this rating was somewhat lower.

Teleguidance was rated as having contributed to cannulation in 17 of 140 (11.9%) cases, and classified as difficult cannulations in 11 of 39 (28%) cases. In 11 of 140 (7.7%) cases, teleguidance was rated as having contributed to stent placement, to stone clearance in 9 of 140 (6.3%) cases, to general assessment in 103 of 140 (72%) cases, and to combinations of these contributions in 13 of 140 (9.3%) cases (Figure 3).

Figure 3. Contributions of teleguidance.



In 23 of 140 cases, surgeons considered teleguidance to have contributed to intervention success and avoiding a repeated attempt at the same intervention (re-ERCP). In 3 of 140 cases, PTC—a more painful and invasive procedure than ERCP—was avoided. In 11 of 140 cases, referral to another ERCP center could be avoided, and in 11 of 140 cases, combinations of the above, and 3 of 140 unspecified other interventions could be avoided (Table 3).

Table 3. Procedures avoided owing to teleguidance (N=140).

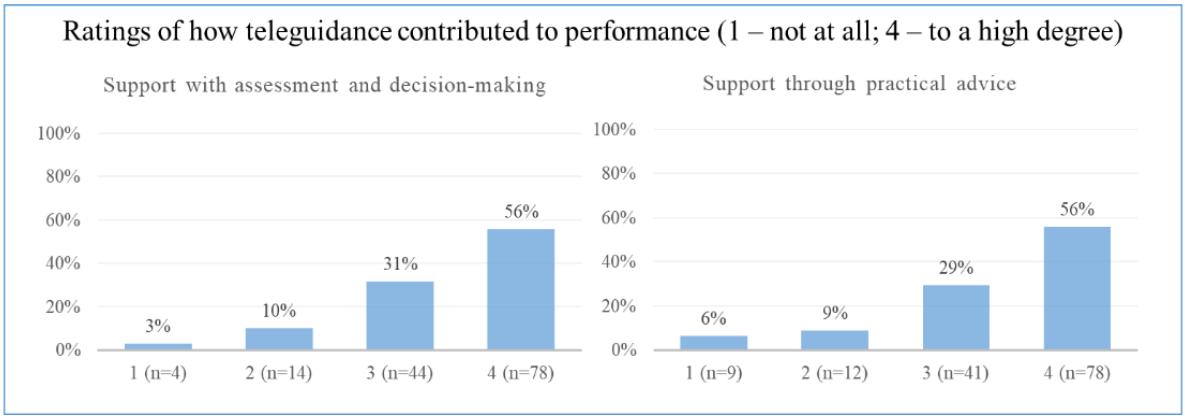
Procedures avoided	Cases, n (%)
Re-ERCP[a]	23 (16.1)
Percutaneous transhepatic cholangiography	3 (2.1)
Other intervention	3 (2.1)
Referral	11 (7.7)
Combinations	11 (7.7)

[a]ERCP: endoscopic retrograde cholangiopancreatography.

In 119 of 140 (85%) cases, teleguidance was reported as having contributed through practical advice to a moderate or large extent, and 122 of 140 (87%) reported that they received support with assessment and decision-making during the procedure to a moderate or large extent (Figure 4).

Overall, the satisfaction ratings, which were measured after procedures, were higher than perceived usefulness ratings, which were measured prior to procedures.

Figure 4. Teleguidance provided support through practical advice and help with assessment.



Discussion

When a telemedicine service for intraoperative surgical consultation was scaled up, we were interested in users' perceptions of how the service contributed to performance and outcomes, as this might provide insight into adoption and use of the telemedicine service over time. We designed and collected measures of perceived usefulness and satisfaction in direct conjunction with real-time use in ERCP procedures. The measures were intended to reflect how users considered teleguidance to contribute to performance and outcomes during ERCP procedures.

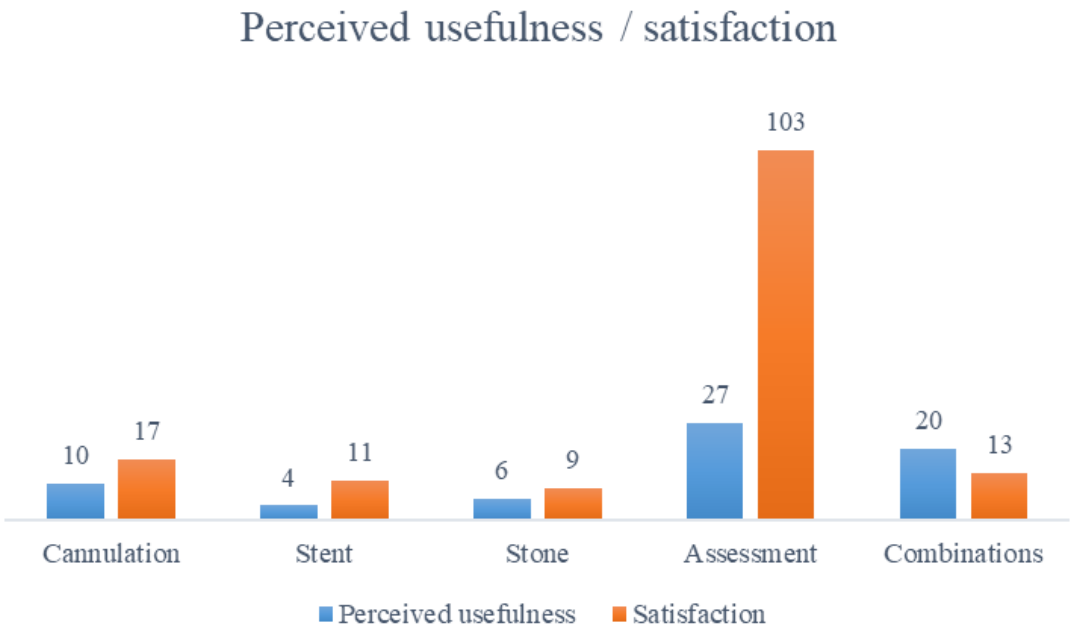
Practitioners believed that teleguidance would be useful; that is, as having value for performance and outcomes, prior to a high proportion of cases. In roughly half of the cases, surgeons specified the type of support they expected, which, in most cases, was related to clinical assessment (27/67, 40.3%). However, the anticipated demand for teleguidance increased with the level of procedural complexity (Multimedia Appendix 14), and there was more interest for teleguidance in certain clinical indications, such as acute and chronic pancreatitis, PSC, and strictures of unknown type (Multimedia Appendix 11).

The results showed that less experienced practitioners perceived teleguidance as more useful than their experienced colleagues did, but the findings were not significant. In addition, the perceived usefulness of teleguidance was higher in cases that could be expected to be challenging.

Regarding satisfaction with teleguidance after procedures, the operating endoscopist rated teleguidance to have contributed to performance and outcomes to a moderate or large extent in 111 of 140 (79.3%) cases. Specific examples are as follows: contribution to intervention success and avoiding a repeated ERCP in 23 cases, avoiding 3 PTC, and 11 referrals, and in 11 cases, combinations of these outcomes (Multimedia Appendix 15).

Our results show that satisfaction after using teleguidance was higher than the preprocedure usefulness of teleguidance, which was rated prior to the procedures (Figure 5). This indicates that it is difficult for practitioners to predict how a novel way of working, such as teleguidance, can contribute to performance and outcomes. User beliefs and attitudes toward technology can be expected to change with first-hand use [55]. Our findings indicate that doctors may become more cognizant of how remote surgical consultation can support important clinical and development/training aspects in ERCP [3] with hands-on use, but also that they require some time before they assimilate teleguidance into their practice. The technical issues experienced did cause some inconvenience and delays, but did not appear to cause teleguidance sessions to be terminated. However, the satisfaction ratings were lower in the cases where technical issues were encountered.

Figure 5. Our results show that satisfaction after using teleguidance was higher than beliefs about usefulness prior to procedures.



While training and assessment of surgical performance is commonly focused on technical ability, cognitive and social skills are also important requirements for surgical competence. Teleguidance may be of value for these nontechnical skills for surgeons, which have been defined as "behavioral aspects of performance in the operating room which underpin medical expertise, use of equipment and drugs: cognitive (e.g. situation awareness, decision making) social (e.g. communication & teamwork, leadership) skills" [56].

Research has shown that user beliefs and attitudes toward technology can be expected to change with first-hand use [55]: experience of teleguidance may gradually change over time. This study also only focused on clinicians' experience of teleguidance during procedures and does not consider contextual factors that are believed to affect the acceptance of teleguidance

[3]. Therefore, these results represent an interim judgement of usefulness and satisfaction, which may differ from final overall satisfaction outcomes [35].

Conclusions

Surgeons appeared not to have expected the level of support they received through remote surgical consultation during ERCP. They also received help with surgical/technical tasks, such as stent placement and stone removal. Each case of support may be of high value from a patient's perspective and for ERCP quality and health economic reasons. The difference between preprocedure expectations and postprocedure satisfaction indicates that practitioners require hands-on use experience to understand the usefulness of the new telemedicine service and how it contributes to surgical procedures. For this reason, adoption can be expected to develop over time and require extended use before being accepted.

While a larger sample of procedures is required to be able to draw statistical inferences about the contribution of teleguidance on clinical outcomes, the findings from the survey items on perceived usefulness and satisfaction indicate that surgeons consider teleguidance to contribute to nontechnical aspects of surgical performance, such as decision-making, to an extent that many practitioners did not anticipate.

This study represents part of a human-centered approach to system design, where system quality is linked to how well users can achieve specific goals. From a methodological perspective, it would be interesting to investigate how interim measures of acceptance and satisfaction correspond with final acceptance and use of the telemedicine service, which appears to be a lesser investigated area [35]. In the case of teleguidance, the service was largely abandoned after the initial start-up phase, even though it is intermittently used in-house at the central hospital.

This study also indicates a need to more deeply investigate how remote surgical contributes to clinical procedures, and the ways in which this way of working differs from on-site surgical consultation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Endoscopic retrograde cholangiography (ERCP) complexity rating.
[PNG File , 53 KB - humanfactors_v8i4e30867_app1.png]

Multimedia Appendix 2
Case report form (CRF) items relating to perceived usefulness and satisfaction.
[PNG File , 41 KB - humanfactors_v8i4e30867_app2.png]

Multimedia Appendix 3
A protoype for telestration was tested during development but not included in the solution that was used in this study.
[PNG File , 743 KB - humanfactors_v8i4e30867_app3.png]

Multimedia Appendix 4
Protocol for establishing a connections between the hospitals.
[PDF File (Adobe PDF File), 149 KB - humanfactors_v8i4e30867_app4.pdf]

Multimedia Appendix 5
Indications and aims for the procedures.
[PNG File , 33 KB - humanfactors_v8i4e30867_app5.png]

Multimedia Appendix 6
Patient sex and age, and case complexity ratings.
[PNG File , 13 KB - humanfactors_v8i4e30867_app6.png]

Multimedia Appendix 7
Distribution of cases across hospitals and practitioners.
[PNG File , 30 KB - humanfactors_v8i4e30867_app7.png]

Multimedia Appendix 8
The operating surgeon's level of expertise and the demand for teleguidance.
[PNG File , 18 KB - humanfactors_v8i4e30867_app8.png]

Multimedia Appendix 9
Intervention goals and anticipated demand for teleguidance.
[PNG File , 54 KB - humanfactors_v8i4e30867_app9.png]

Multimedia Appendix 10
Anticipated demand for teleguidance for different indications.
[PNG File , 72 KB - humanfactors_v8i4e30867_app10.png]

Multimedia Appendix 11
Situations which surgeons believed teleguidance might help them might avoid.
[PNG File , 16 KB - humanfactors_v8i4e30867_app11.png]

Multimedia Appendix 12
Tasks that surgeons expected to get support with through teleguidance.
[PNG File , 19 KB - humanfactors_v8i4e30867_app12.png]

Multimedia Appendix 13
Ratings of the value of teleguidance.
[PNG File , 11 KB - humanfactors_v8i4e30867_app13.png]

Multimedia Appendix 14
A higher level of case complexity corresponded with a higher demand for teleguidance.
[PNG File , 25 KB - humanfactors_v8i4e30867_app14.png]

Multimedia Appendix 15
In many cases, practitioners considered teleguidance to have have helped avoid unwanted outcomes.
[PNG File , 11 KB - humanfactors_v8i4e30867_app15.png]

References

1. Påhlsson HI, Groth K, Permert J, Swahn F, Löhr M, Enochsson L, et al. Telemedicine: an important aid to perform high-quality endoscopic retrograde cholangiopancreatography in low-volume centers. Endoscopy 2013;45(5):357-361. [doi: 10.1055/s-0032-1326269] [Medline: 23468194]
2. Brinne Roos J, Bergenzaun P, Groth K, Lundell L, Arnelo U. Telepresence-teleguidance to facilitate training and quality assurance in ERCP: a health economic modeling approach. Endosc Int Open 2020 Mar;8(3):E326-E337 [FREE Full text] [doi: 10.1055/a-1068-9153] [Medline: 32118106]
3. Aminoff H, Meijer S, Arnelo U, Groth K. Modeling the implementation context of a telemedicine service: work domain analysis in a surgical setting. JMIR Form Res 2021 Jun 21;5(6):e26505 [FREE Full text] [doi: 10.2196/26505] [Medline: 34152278]
4. Aminoff H, Meijer S, Arnelo U, Frennert S. Telemedicine for remote surgical guidance in endoscopic retrograde cholangiopancreatography: mixed methods study of practitioner attitudes. JMIR Form Res 2021 Jan 11;5(1):e20692 [FREE Full text] [doi: 10.2196/20692] [Medline: 33427670]
5. Groth K, Arnelo U, Bergenzaun P, Lundell L, Aminoff H, Frykholm O, et al. Designing for interactivity in a tele-guidance setting. Int J Integr Care 2016 Nov 09;16(5):19. [doi: 10.5334/ijic.2568]
6. Langerth A, Isaksson B, Karlson B, Urdzik J, Linder S. ERCP-related perforations: a population-based study of incidence, mortality, and risk factors. Surg Endosc 2020 May;34(5):1939-1947 [FREE Full text] [doi: 10.1007/s00464-019-06966-w] [Medline: 31559577]
7. Enochsson L, Swahn F, Arnelo U, Nilsson M, Löhr M, Persson G. Nationwide, population-based data from 11,074 ERCP procedures from the Swedish Registry for Gallstone Surgery and ERCP. Gastrointest Endosc 2010 Dec;72(6):1175-84, 1184.e1. [doi: 10.1016/j.gie.2010.07.047] [Medline: 20970787]
8. Keswani RN, Qumseya BJ, O'Dwyer LC, Wani S. Association between endoscopist and center endoscopic retrograde cholangiopancreatography volume with procedure success and adverse outcomes: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2017 Dec;15(12):1866-1875.e3. [doi: 10.1016/j.cgh.2017.06.002] [Medline: 28606848]
9. Rutter M, Senore C, Bisschops R, Domagk D, Valori R, Kaminski M, et al. The European Society of Gastrointestinal Endoscopy Quality Improvement Initiative: developing performance measures. Endoscopy 2016 Jan;48(1):81-89 [FREE Full text] [doi: 10.1055/s-0035-1569580] [Medline: 26662057]

10. Noel R, Enochsson L, Swahn F, Löhr M, Nilsson M, Permert J, et al. A 10-year study of rendezvous intraoperative endoscopic retrograde cholangiography during cholecystectomy and the risk of post-ERCP pancreatitis. Surg Endosc 2013 Jul;27(7):2498-2503. [doi: 10.1007/s00464-012-2768-4] [Medline: 23355164]
11. Bailey A, Bourke M, Williams S, Walsh P, Murray M, Lee E, et al. A prospective randomized trial of cannulation technique in ERCP: effects on technical success and post-ERCP pancreatitis. Endoscopy 2008 Apr;40(4):296-301. [doi: 10.1055/s-2007-995566] [Medline: 18389448]
12. Adler DG, Lieb JG, Cohen J, Pike IM, Park WG, Rizk MK, et al. Quality indicators for ERCP. Gastrointest Endosc 2015 Jan;81(1):54-66. [doi: 10.1016/j.gie.2014.07.056] [Medline: 25480099]
13. Lee HJ, Cho CM, Heo J, Jung MK, Kim TN, Kim KH, et al. Impact of Hospital Volume and the Experience of Endoscopist on Adverse Events Related to Endoscopic Retrograde Cholangiopancreatography: A Prospective Observational Study. Gut Liver 2020 Mar 15;14(2):257-264 [FREE Full text] [doi: 10.5009/gnl18537] [Medline: 31060119]
14. Augestad KM, Lindsetmo RO. Overcoming distance: video-conferencing as a clinical and educational tool among surgeons. World J Surg 2009 Jul;33(7):1356-1365 [FREE Full text] [doi: 10.1007/s00268-009-0036-0] [Medline: 19384459]
15. Augestad KM, Han H, Paige J, Ponsky T, Schlachta CM, Dunkin B, et al. Educational implications for surgical telementoring: a current review with recommendations for future practice, policy, and research. Surg Endosc 2017 Oct;31(10):3836-3846. [doi: 10.1007/s00464-017-5690-y] [Medline: 28656341]
16. Erridge S, Yeung DKT, Patel HRH, Purkayastha S. Telementoring of surgeons: a systematic review. Surg Innov 2019 Feb;26(1):95-111. [doi: 10.1177/1553350618813250] [Medline: 30465477]
17. Cubano M, Poulose BK, Talamini MA, Stewart R, Antosek LE, Lentz R, et al. Long distance telementoring. A novel tool for laparoscopy aboard the USS Abraham Lincoln. Surg Endosc 1999 Jul;13(7):673-678. [doi: 10.1007/s004649901071] [Medline: 10384073]
18. Taniguchi E, Ohashi S. Construction of a regional telementoring network for endoscopic surgery in Japan. IEEE Trans Inf Technol Biomed 2000 Sep;4(3):195-199. [doi: 10.1109/4233.870029] [Medline: 11026589]
19. Zanaboni P, Wootton R. Adoption of telemedicine: from pilot stage to routine delivery. BMC Med Inform Decis Mak 2012 Jan 04;12:1 [FREE Full text] [doi: 10.1186/1472-6947-12-1] [Medline: 22217121]
20. Antoniou SA, Antoniou GA, Franzen J, Bollmann S, Koch OO, Pointner R, et al. A comprehensive review of telementoring applications in laparoscopic general surgery. Surg Endosc 2012 Aug;26(8):2111-2116. [doi: 10.1007/s00464-012-2175-x] [Medline: 22350150]
21. Prabhakaran K, Lombardo G, Latifi R. Telemedicine for trauma and emergency management: an overview. Curr Trauma Rep 2016 May 27;2(3):115-123. [doi: 10.1007/s40719-016-0050-2]
22. Huang EY, Knight S, Guetter CR, Davis CH, Moller M, Slama E, et al. Telemedicine and telementoring in the surgical specialties: A narrative review. Am J Surg 2019 Oct;218(4):760-766. [doi: 10.1016/j.amjsurg.2019.07.018] [Medline: 31350010]
23. Armfield NR, Edirippulige SK, Bradford N, Smith AC. Telemedicine--is the cart being put before the horse? Med J Aust 2014 May 19;200(9):530-533. [doi: 10.5694/mja13.11101] [Medline: 24835716]
24. Budrionis A, Augestad KM, Patel HR, Bellika JG. An evaluation framework for defining the contributions of telestration in surgical telementoring. Interact J Med Res 2013 Jul 25;2(2):e14 [FREE Full text] [doi: 10.2196/ijmr.2611] [Medline: 23887078]
25. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
26. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
27. Rubinger L, Bhandari M. Cochrane in CORR®-Interactive Telemedicine: effects on professional practice and health care outcomes. Clin Orthop Relat Res 2020 Sep;478(9):1971-1973 [FREE Full text] [doi: 10.1097/CORR.0000000000001440] [Medline: 32769532]
28. Agboola S, Hale TM, Masters C, Kvedar J, Jethwani K. "Real-world" practical evaluation strategies: a review of telehealth evaluation. JMIR Res Protoc 2014 Dec 17;3(4):e75 [FREE Full text] [doi: 10.2196/resprot.3459] [Medline: 25524892]
29. Kidholm K, Clemensen J, Caffery LJ, Smith AC. The Model for Assessment of Telemedicine (MAST): a scoping review of empirical studies. J Telemed Telecare 2017 Oct;23(9):803-813. [doi: 10.1177/1357633X17721815] [Medline: 28758525]
30. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
31. Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. Int J Med Inform 2010 Nov;79(11):736-771. [doi: 10.1016/j.ijmedinf.2010.08.006] [Medline: 20884286]
32. ISO 9241-11:2018 Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. International Organization for Standardization. URL: https://www.iso.org/standard/63500.html [accessed 2021-11-09]
33. Lewis J, Sauro J. The Factor Structure of the System Usability Scale. In: Human Centered Design. Berlin, Heidelberg: Springer; 2009:94-103.

34. Wixom BH, Todd PA. A theoretical integration of user satisfaction and technology acceptance. Inf Syst Res 2005 Mar;16(1):85-102. [doi: 10.1287/isre.1050.0042]
35. Vaezi R, Mills A, Chin W, Zafar H. User satisfaction research in information systems: historical roots and approaches. CAIS 2016;38:501-532. [doi: 10.17705/1cais.03827]
36. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
37. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man-Machine Studies 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
38. Garcia R, Adelakun O. Towards a Conceptual Framework for the Evaluation of Telemedicine Satisfaction. In: Trends and Advances in Information Systems and Technologies. Cham: Springer; 2018:1188-1197.
39. Kidholm K, Ekeland AG, Jensen LK, Rasmussen J, Pedersen CD, Bowes A, et al. A model for assessment of telemedicine applications: mast. Int J Technol Assess Health Care 2012 Jan;28(1):44-51. [doi: 10.1017/S0266462311000638] [Medline: 22617736]
40. Whitten P, Love B. Patient and provider satisfaction with the use of telemedicine: overview and rationale for cautious enthusiasm. J Postgrad Med 2005;51(4):294-300 [FREE Full text] [Medline: 16388172]
41. Ekeland AG, Bowes A, Flottorp S. Methodologies for assessing telemedicine: a systematic review of reviews. Int J Med Inform 2012 Jan;81(1):1-11. [doi: 10.1016/j.ijmedinf.2011.10.009] [Medline: 22104370]
42. Garcia R, Adelakun O. A conceptual framework and pilot study for examining telemedicine satisfaction research. J Med Syst 2019 Jan 25;43(3):51. [doi: 10.1007/s10916-019-1161-4] [Medline: 30684065]
43. Hirani SP, Rixon L, Beynon M, Cartwright M, Cleanthous S, Selva A, et al. Quantifying beliefs regarding telehealth: Development of the Whole Systems Demonstrator Service User Technology Acceptability Questionnaire. J Telemed Telecare 2017 May;23(4):460-469. [doi: 10.1177/1357633X16649531] [Medline: 27224997]
44. Nguyen M, Waller M, Pandya A, Portnoy J. A review of patient and provider satisfaction with telemedicine. Curr Allergy Asthma Rep 2020 Sep 22;20(11):72 [FREE Full text] [doi: 10.1007/s11882-020-00969-7] [Medline: 32959158]
45. Hajesmaeel-Gohari S, Bahaadinbeigy K. The most used questionnaires for evaluating telemedicine services. BMC Med Inform Decis Mak 2021 Feb 02;21(1):36 [FREE Full text] [doi: 10.1186/s12911-021-01407-y] [Medline: 33531013]
46. AlDossary S, Martin-Khan MG, Bradford NK, Smith AC. A systematic review of the methodologies used to evaluate telemedicine service initiatives in hospital facilities. Int J Med Inform 2017 Jan;97:171-194. [doi: 10.1016/j.ijmedinf.2016.10.012] [Medline: 27919377]
47. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015 Sep 07(9):CD002098 [FREE Full text] [doi: 10.1002/14651858.CD002098.pub2] [Medline: 26343551]
48. World Health Organization. Medical Devices: Managing the Mismatch : an Outcome of the Priority Medical Devices Project. Albany, NY: World Health Organization; 2010.
49. Standing C, Standing S, McDermott M, Gururajan R, Kiani Mavi R. The paradoxes of telehealth: a review of the literature 2000-2015. Syst Res 2016 Dec 01;35(1):90-101. [doi: 10.1002/sres.2442]
50. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
51. Cho S, Mathiassen L, Gallivan M. Crossing the diffusion chasm: from invention to penetration of a telehealth innovation. Info Technology & People 2009 Nov 13;22(4):351-366. [doi: 10.1108/09593840911002450]
52. Testoni P, Mariani A, Aabakken L, Arvanitakis M, Bories E, Costamagna G, et al. Papillary cannulation and sphincterotomy techniques at ERCP: European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. Endoscopy 2016 Jul;48(7):657-683 [FREE Full text] [doi: 10.1055/s-0042-108641] [Medline: 27299638]
53. Haraldsson E, Lundell L, Swahn F, Enochsson L, Löhr JM, Arnelo U, Scandinavian Association for Digestive Endoscopy (SADE) Study Group of Endoscopic Retrograde Cholangio-Pancreaticography. Endoscopic classification of the papilla of Vater. Results of an inter- and intraobserver agreement study. United European Gastroenterol J 2017 Jun;5(4):504-510 [FREE Full text] [doi: 10.1177/2050640616674837] [Medline: 28588881]
54. Budrionis A, Hasvold P, Hartvigsen G, Bellika JG. Assessing the impact of telestration on surgical telementoring: a randomized controlled trial. J Telemed Telecare 2016 Jan;22(1):12-17. [doi: 10.1177/1357633X15585071] [Medline: 26026177]
55. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
56. Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non-technical skills. Med Educ 2006 Nov;40(11):1098-1104. [doi: 10.1111/j.1365-2929.2006.02610.x] [Medline: 17054619]

Abbreviations

CRF: case report form

ERC: endoscopic retrograde cholangiography
ERCP: endoscopic retrograde cholangiopancreatography
PSC: primary sclerosing cholangitis
PTC: percutaneous transhepatic cholangiography
TAM: technology acceptance model

Edited by A Kushniruk; submitted 02.06.21; peer-reviewed by K Augestad, À Salvador Verges; comments to author 04.07.21; revised version received 16.07.21; accepted 23.07.21; published 30.11.21.

Please cite as:
Aminoff H, Meijer S, Groth K, Arnelo U
User Experience in Remote Surgical Consultation: Survey Study of User Acceptance and Satisfaction in Real-Time Use of a Telemedicine Service
JMIR Hum Factors 2021;8(4):e30867
URL: https://humanfactors.jmir.org/2021/4/e30867
doi:10.2196/30867
PMID:34851302

Original Paper

Clinician Preimplementation Perspectives of a Decision-Support Tool for the Prediction of Cardiac Arrhythmia Based on Machine Learning: Near-Live Feasibility and Qualitative Study

Stina Matthiesen[1,2], PhD; Søren Zöga Diederichsen[2,3], MD; Mikkel Klitzing Hartmann Hansen[2], PhD; Christina Villumsen[2], MSc; Mats Christian Højbjerg Lassen[2], BSc; Peter Karl Jacobsen[3], DMSc, MD; Niels Risum[3], MD, PhD; Bo Gregers Winkel[3], MD, PhD; Berit T Philbert[3], MD, PhD; Jesper Hastrup Svendsen[3,4], DMSc, MD; Tariq Osman Andersen[1,2], PhD

[1]Department of Computer Science, Faculty of Science, University of Copenhagen, Copenhagen, Denmark

[2]Vital Beats, Copenhagen, Denmark

[3]Department of Cardiology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

[4]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Stina Matthiesen, PhD
Department of Computer Science
Faculty of Science
University of Copenhagen
Universitetsparken 5
Copenhagen, 2100
Denmark
Phone: 45 21231008
Email: matthiesen@di.ku.dk

Abstract

Background: Artificial intelligence (AI), such as machine learning (ML), shows great promise for improving clinical decision-making in cardiac diseases by outperforming statistical-based models. However, few AI-based tools have been implemented in cardiology clinics because of the sociotechnical challenges during transitioning from algorithm development to real-world implementation.

Objective: This study explored how an ML-based tool for predicting ventricular tachycardia and ventricular fibrillation (VT/VF) could support clinical decision-making in the remote monitoring of patients with an implantable cardioverter defibrillator (ICD).

Methods: Seven experienced electrophysiologists participated in a near-live feasibility and qualitative study, which included walkthroughs of 5 blinded retrospective patient cases, use of the prediction tool, and questionnaires and interview questions. All sessions were video recorded, and sessions evaluating the prediction tool were transcribed verbatim. Data were analyzed through an inductive qualitative approach based on grounded theory.

Results: The prediction tool was found to have potential for supporting decision-making in ICD remote monitoring by providing reassurance, increasing confidence, acting as a second opinion, reducing information search time, and enabling delegation of decisions to nurses and technicians. However, the prediction tool did not lead to changes in clinical action and was found less useful in cases where the quality of data was poor or when VT/VF predictions were found to be irrelevant for evaluating the patient.

Conclusions: When transitioning from AI development to testing its feasibility for clinical implementation, we need to consider the following: expectations must be aligned with the intended use of AI; trust in the prediction tool is likely to emerge from real-world use; and AI accuracy is relational and dependent on available information and local workflows. Addressing the sociotechnical gap between the development and implementation of clinical decision-support tools based on ML in cardiac care is essential for succeeding with adoption. It is suggested to include clinical end-users, clinical contexts, and workflows throughout the overall iterative approach to design, development, and implementation.

(JMIR Hum Factors 2021;8(4):e26964) doi:10.2196/26964

KEYWORDS

cardiac arrhythmia; short-term prediction; clinical decision support systems; machine learning; artificial intelligence; preimplementation; qualitative study; implantable cardioverter defibrillator; remote follow-up; sociotechnical

Introduction

Ventricular tachycardia and ventricular fibrillation (VT/VF) are potentially lethal cardiac arrhythmias, which constitute a growing challenge to health care systems worldwide [1]. The development of implantable cardioverter defibrillators (ICDs) has led to major advances in the prevention of death from VT/VF [2]. ICDs are implantable devices used in patients at increased risk of sudden cardiac death. ICDs monitor the heart rhythm continuously to detect and treat VT/VF. In recent years, remote monitoring has become the standard of care for ICD patients [3], and follow-ups are based on transmission of data from the implanted device through the patient's home monitoring box. This has reduced the number of in-office follow-ups [4,5] and increased survival rates [6] due to improved early detection of arrhythmias [7]. However, the numbers of ICD implants are increasing worldwide, posing a workload challenge for electrophysiologists and technicians when assessing data from incoming transmissions in remote monitoring centers [8-11]. There is a growing need for decision-making tools that can support and reduce data-intensive remote follow-ups, and while current systems can detect and treat VT/VF arrhythmias as they occur, tools for predicting arrhythmias before their onset are lacking [12].

Artificial intelligence (AI), such as machine learning (ML), shows great promise for improving clinical decision-making in cardiac diseases by outperforming statistical-based models [12,13], and recent examples include promising models for the prediction of heart disease and heart failure [14-18], as well as cardiac arrhythmias, such as ventricular arrhythmia [19], atrial fibrillation [20], and electrical storm [21]. There are positive attitudes and high expectations among physicians that AI will improve future patient care in fields where data are collected continuously, such as cardiology [22,23].

However, few prediction outcome algorithms based on ML have been implemented in cardiology clinics because of the challenges during transitioning from algorithm development to real-world implementation. While studies of medical AI-based tools that undergo prospective clinical validation are emerging [24-27], there is a general lack of understanding of how AI may support achieving clinical effectiveness and improve patient care in real-life settings [28,29]. Scholars have argued that ML-based patient outcome prediction models are yet to prove their worth to human clinicians [30]. Prediction accuracy by itself can be impressive in the lab; however, this does not always translate to better treatment, and it is being stressed to look for ways to make human and AI prediction algorithms complement each other, ensuring actionability in clinical practice [30-33]. Going from research and development environments to hospital or clinical contexts is considered a challenging task that has been named "the last mile" of implementing medical AI-based tools [34,35], and there is a call for research on how end-users find AI-based user interfaces useful in practice [36-39], as well as studies that report on the sociotechnical challenges of deploying AI-based tools in complex clinical environments [27,34,35,40-54].

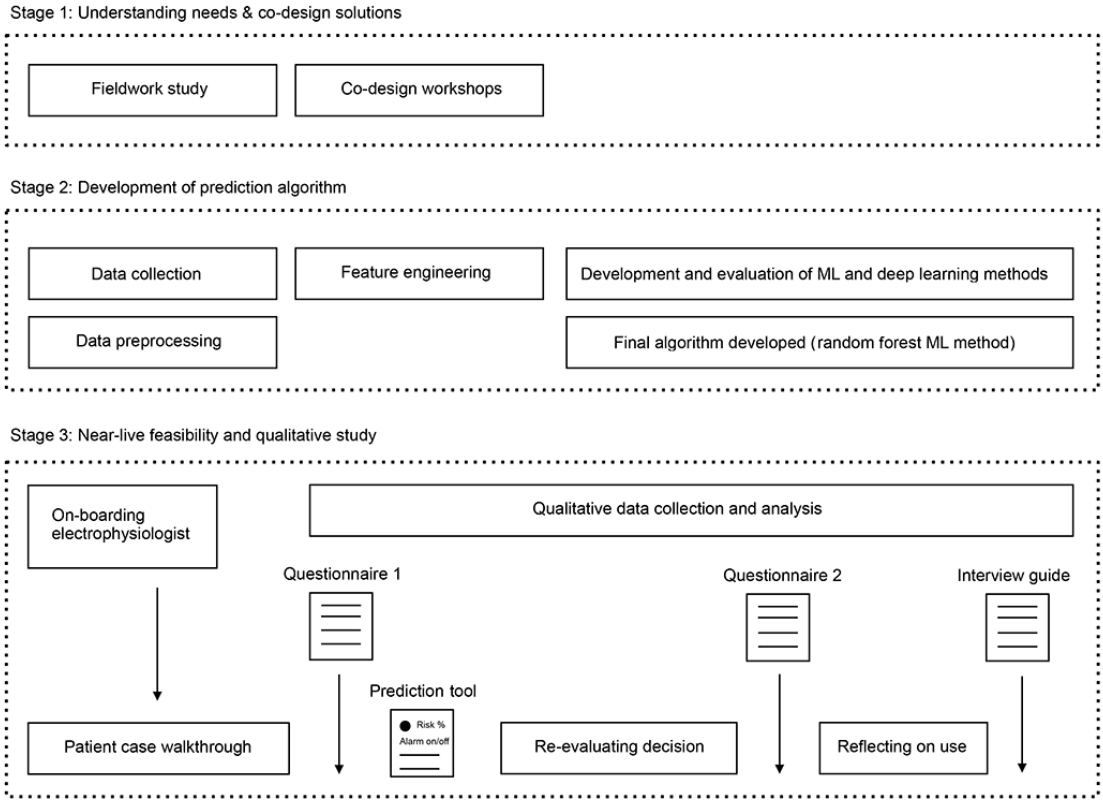
This study addresses the sociotechnical gap between the development and implementation of a clinical decision-support tool based on ML for the prediction of VT/VF in remote monitoring of ICD patients. The aim of this study was to explore the feasibility and clinician preimplementation perspectives of using a prediction tool for improved workflows. Therefore, this study does not provide algorithmic validation per se but instead answers questions about the clinical feasibility and workflow integration of a decision-support tool based on ML.

Methods

Understanding Needs and Co-design of the Prediction Tool

This study was conducted at the remote monitoring center at Rigshospitalet, Copenhagen University Hospital, Denmark, which is a large tertiary hospital covering all aspects of treatments in cardiology and is among the largest centers in Europe having more than 4000 patients with cardiac implanted electronic devices in remote follow-up. The study was organized in 3 stages (Figure 1). In the first stage, field work observations in the remote monitoring clinic were conducted to understand both the clinical workflow and workload [10]. This was followed by 3 co-design workshops with an electrophysiologist (PKJ) and 5 co-design workshops with a cardiologist consultant (SZD) focusing on feature engineering and sketching the user interface. In stage 2, the AI algorithm was developed, and in stage 3, a near-live feasibility and qualitative interview study was conducted. The study was reviewed by the Danish National Board of Health and the Danish National Committee on Health Research Ethics, and authorized by The Capital Region of Denmark.

Figure 1. Overall study design. ML: machine learning.



Development of the AI Algorithm

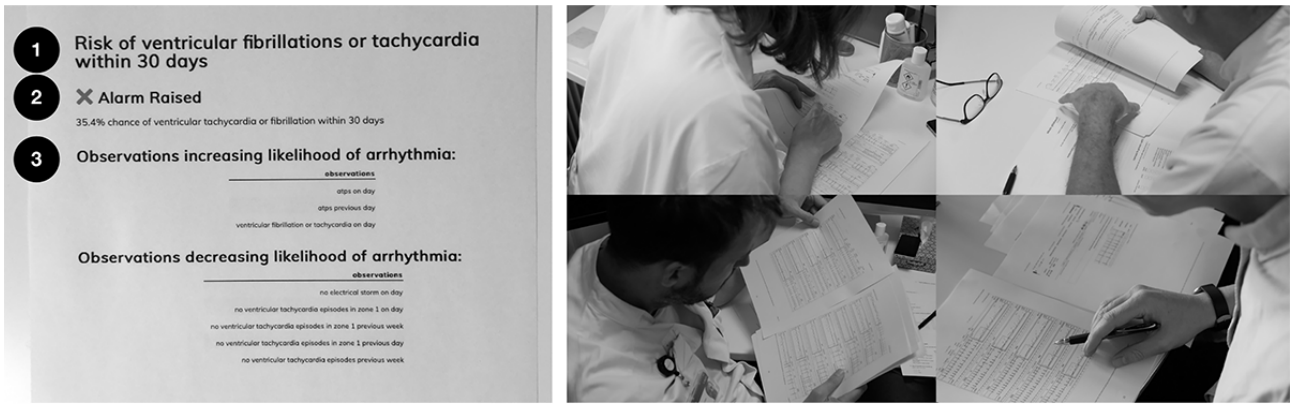
A prediction tool was developed for improving the support for clinical decision-making in ICD remote monitoring based on the random forest ML method, and it consisted of a risk prediction algorithm of VT/VF within 30 days. The prediction tool was designed to show alarm status (yes/no), risk probability (%), and ranking of the 5 most and least important parameters for the prediction, using the LIME technique [55] (Figure 2). The design and development of the tool were informed by previous fieldwork studies of current practices [10,56], as well as early results from using ML to predict electrical storm, a severe form of cardiac arrhythmia [21]. The data set used for developing the algorithm consisted of 11,921 transmissions from 1251 patients with an ICD or a cardiac resynchronization therapy defibrillator (CRT-D), followed over a 4-year period from 2015 to 2019 at Rigshospitalet. The data set contained 74,149 arrhythmia episodes, each characterized by 7 variables, such as the type of arrhythmia (VT, VF, supraventricular tachycardia, atrial fibrillation, etc), ICD treatment of the arrhythmia, duration of the episode, and maximum heart rate reached during the episode.

The random forest ML method [44] was selected for algorithm development because it provided optimal results when considering the tradeoffs between model performance and explainability. Several other classifier methods (supervised, unsupervised, and deep learning methods) were evaluated through development and testing, including KNeighborsClassifier [57], GradientBoostingClassifier [58], AdaBoostClassifier [59], support vector classifier [60], and long short-term memory (LSTM) [61]. The deep learning method, LSTM, provided poorer performance and poorer explainability, possibly due to the nature of the data (ie, time series data with considerable time between events, making time series modeling difficult). The other methods provided similar performance. KNeighborsClassifier and support vector machine had the worst performance, while the decision tree methods had the best performance. GradientBoostingClassifier produced an optimal F1 score and recall score; however, random forest provided the highest accuracy and precision scores, which led to the choice of using the random forest method for developing the first version of the algorithm to be evaluated with end-users in this study. The algorithm was tested on 2342 of the 11,921 transmissions. The transmission data were stratified and grouped into training and test sets. This means that the prevalence of the positive condition was the same in both the training and test sets (stratified) and that no patient had data in both data sets (grouped). The algorithm achieved an accuracy of 0.96, with a positive predictive value of 0.67 and a negative predictive value of 0.97. The probability threshold for raising an alarm was set to 0.28, indicating the value with an optimal tradeoff between negative and positive predictive outcomes.

Feature engineering was carried out in collaboration between 2 data scientists (MKHH and CV) and a cardiologist consultant (SZD) during 5 co-design workshops. A total of 48 features (referred to as parameters when discussed with the study participants) were developed, and the following 2 main principles were adopted: aggregating episodes by day and building a historic snapshot for days leading up to the arrhythmic event. To provide the clinical end-user with algorithm explainability, the LIME technique [55] was used to show the

top 5 features that increase or decrease the likelihood of a VT/VF arrhythmic event occurring within the coming 30 days.

Figure 2. The prediction tool on a paper printout as shown to study participants (Case 3, see Table 2). The output shows the alarm (yes/no), risk probability (%), and up to 5 most important parameters for increasing and decreasing the likelihood of ventricular tachycardia and ventricular fibrillation within 30 days. To the right: example pictures of electrophysiologists conducting near-live case walkthroughs.



Study Participants and Case Selection

Seven medical doctors specialized in electrophysiology (ie, cardiologists treating patients with cardiac arrhythmia) were selected for participation from a convenience sample (Table 1). Participants included 6 males and 1 female (average age, 52 years; average work experience as an electrophysiologist, 13 years).

A selection of 5 retrospective patient cases (Table 2) was used to evaluate the feasibility of the prediction tool's ability to support clinical decision-making. The cases included high and low risk probability, true positives, and true negatives, and 2 cases with AF as the primary episode type. Patient cases were retrieved 19 to 27 months back in time, blinded, and presented as paper printouts with a summary of each patient's clinical history along with reports from the electronic health record (list of diagnoses, progress notes from the cardiology department, latest blood tests, and list of medications) and screenshots of relevant ICD transmission data, including device type, battery status, device programming and settings, time of implantation, latest diagnostic information about the transmission, frequency of arrhythmias, heart rate, device therapy, and assessment of physical activity.

Table 1. Participating electrophysiologists.

Participant	Sex	Age (years)	Title	Years since obtaining specialist certification in cardiology
1	Female	52	Consultant cardiologist, MD, PhD	11
2	Male	61	Professor, consultant cardiologist, MD, DMSc	23
3	Male	55	Consultant cardiologist, MD, PhD	14
4	Male	43	Cardiologist, MD, PhD	2
5	Male	62	Consultant cardiologist, MD, DMSc	28
6	Male	44	Cardiologist, MD, PhD	2
7	Male	47	Consultant cardiologist, MD, DMSc	9

Table 2. Case overview with patient summary, current implantable cardioverter defibrillator transmission information, and prediction tool information.

Case number	Patient summary	Current ICD[a] transmission				Prediction tool	
		Transmission type	Primary episode type	ICD treatment	Transmission summary	30-day VT[b]/VF[c] risk probability	Alarm raised (prediction outcome)
1	Male, age 63 years, ischemic heart failure, left ventricular assist device	Automated	VT/VF	ATP[d]	3 VT/VF; 36 sensing episodes; 217 VT-NS[e]	58.6	Yes (true positive)
2	Female, age 67 years, dilated cardiomyopathy	Automated	VT/VF	Shock	1 VT/VF; 1 VT-NS; 20 min of AF[f] since the last transmission	14.4	No (true negative)
3	Female, age 40 years, dilated cardiomyopathy	Automated	VT/VF	Shock	2 VT/VF; 4 VT-NS	35.4	Yes (true positive)
4	Male, age 61 years, ischemic heart failure	Patient initiated	AF	None	12 hours of AF since the last transmission	1.2	No (true negative)
5	Male, age 73 years, ischemic heart failure	Automated	AF	None	14 hours of AF since the last session; 26 VT-NS	7.8	No (true negative)

[a]ICD: implantable cardioverter defibrillator.

[b]VT: ventricular tachycardia.

[c]VF: ventricular fibrillation.

[d]ATP: antitachycardia pacing.

[e]VT-NS: nonsustained ventricular tachycardia.

[f]AF: atrial fibrillation.

Data Collection

A combined feasibility and qualitative interview study was undertaken based on a retrospective case study design. The primary aim of the study was to address the following 4 main questions about the feasibility of the prediction tool using quantitative measures: Does use of the tool lead to change in clinical action? Does it support decision-making? Are visualizing parameters useful? Can it reduce time spent? The secondary aims were to understand the electrophysiologist's immediate reactions to using the prediction tool, including qualifying the quantitative feasibility measures against qualitative dimensions based on interviews. Electrophysiologists were invited to conduct a "near-live" clinical simulation of decision-making based on walkthroughs of the 5 patient cases (Table 2) with and without the prediction tool. Two structured questionnaires based on a 5-point Likert scale were designed to capture electrophysiologists' decisions on action without the prediction tool (Multimedia Appendix 1) and their experiences of the feasibility of the prediction tool (Multimedia Appendix 2). A semistructured interview guide was designed based on the framework of Bowen et al for feasibility studies [62] to cover open-ended questions about the electrophysiologists' overall experiences of using the prediction tool. Ten questions in the following 4 areas of inquiry were posed: acceptability, demand, adoption, and implementation (Multimedia Appendix 3).

"Near-live" case walkthroughs were performed with inspiration from Li et al [63] (Figure 1), and they were facilitated by the authors SM, MKHH, and TOA. First, the electrophysiologist was on-boarded with a presentation of the study objectives, the intended use of the prediction tool, the algorithm development (data set and ML model, as well as results), and the outline of the feasibility and qualitative study processes, and time was provided to resolve open questions. Second, the electrophysiologist was provided with a patient case and asked to do a walkthrough of the case material to reach a decision on clinical action, similar to normal clinical practice, and was asked to answer the first questionnaire and explain the reasoning behind the decision on clinical action. Third, the electrophysiologist received the prediction tool on paper and was asked to answer the second questionnaire for evaluation of the effects of the prediction tool and to share his/her immediate reactions. Fourth, after ending all patient case walkthroughs, the electrophysiologist was interviewed about his/her experience of the feasibility of the prediction tool. The total time for observations and interviews was 12.5 hours, with an average of 1 hour 47 minutes per electrophysiologist. Case walkthroughs and interviews were audio and video recorded, and sections with electrophysiologists' responses to the questionnaires and the open-ended interview were transcribed verbatim.

Data Analysis

Data from electrophysiologists' reactions to the interview study were analyzed using an inductive qualitative approach based on grounded theory [64]. A 2-step iterative coding process was applied beginning with line-by-line coding to support initial analytic decisions about the data. Action codes were developed by using gerunds (a noun form of a verb) to make explicit what electrophysiologists were doing during case walkthroughs and what meaning they derived (eg, "being confirmed," "building trust," and "using to prioritize"). This was done to preserve focus on action and situated processes of electrophysiologists' decision-making, and to turn thematic descriptions into

analytical insights in later stages of the analysis. Focused coding was carried out by iteratively sorting and synthesizing line-by-line codes into themes and subthemes related to the research questions and by constructing key insights. This process allowed for comparing and turning frequently reappearing initial codes across large amounts of data, and obtaining more general and analytically incisive findings (eg, "predictions can serve as a second opinion" and "decision-making workload is reduced when trust in the prediction tool is established"). The entire process was carried out iteratively in collaboration between SM and TOA using the qualitative data analysis software NVivo 12 (QSR International).

Results

Feasibility of the Prediction Tool in Clinical Practice

Does the Prediction Tool Change Clinical Decisions?

Overall, the electrophysiologists did not change their decisions on clinical action when presented with the 30-day VT/VF arrhythmia prediction (Table 3). However, several electrophysiologists found that the prediction tool was helpful (Textbox 1, Quote 1) and increased their confidence in their choice of clinical action, and that the predictions could help prioritize patients (Textbox 1, Quote 2) and determine what action to take in relation to the local circumstances at the clinic (Textbox 1, Quote 3).

Table 3. Effect of the prediction tool on electrophysiologists' decision-making.

Question and answer	Total (N=35), n (%)	Case 1 (N=7), n (%)	Case 2 (N=7), n (%)	Case 3 (N=7), n (%)	Case 4 (N=7), n (%)	Case 5 (N=7), n (%)
Q1: The prediction tool made me change my decision on clinical action						
Yes	1 (3)	1 (14)	0 (0)	0 (0)	0 (0)	0 (0)
No	34 (97)	6 (86)	7 (100)	7 (100)	7 (100)	7 (100)
Q1a: I will contact the patient						
Strongly disagree/disagree	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Neither agree nor disagree	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Agree/strongly agree	35 (100)	7 (100)	7 (100)	7 (100)	7 (100)	7 (100)
Q1b: I will book a procedure or reschedule an existing procedure						
Strongly disagree/disagree	3 (9)	0 (0)	1 (14)	1 (14)	0 (0)	1 (14)
Neither agree nor disagree	6 (17)	1 (14)	0 (0)	0 (0)	4 (57)	1 (14)
Agree/strongly agree	26 (74)	6 (86)	6 (86)	6 (86)	3 (43)	5 (71)
Q1c: I will do something else						
Strongly disagree/disagree	19 (54)	5 (71)	3 (43)	5 (71)	2 (29)	4 (57)
Neither agree nor disagree	5 (14)	1 (14)	1 (14)	1 (14)	2 (29)	0 (0)
Agree/strongly agree	11 (31)	1 (14)	3 (43)	1 (14)	3 (43)	3 (43)
Q2: The prediction tool supported my decision-making						
Strongly disagree/disagree	8 (23)	2 (29)	1 (14)	0 (0)	3 (43)	2 (29)
Neither agree nor disagree	4 (11)	1 (14)	0 (0)	1 (14)	1 (14)	1 (14)
Agree/strongly agree	23 (66)	4 (57)	6 (86)	6 (86)	3 (43)	4 (57)
Q3: The prediction tool's visualization of parameters supported my decision making						
Strongly disagree/disagree	11 (31)	3 (43)	1 (14)	2 (29)	3 (43)	2 (29)
Neither agree nor disagree	3 (9)	1 (14)	0 (0)	0 (0)	1 (14)	1 (14)
Agree/strongly agree	21 (60)	3 (43)	6 (86)	5 (71)	3 (43)	4 (57)
Q4: The prediction tool can help me reach a decision faster						
Strongly disagree/disagree	13 (37)	4 (57)	2 (29)	2 (29)	2 (29)	3 (43)
Neither agree nor disagree	4 (11)	1 (14)	0 (0)	0 (0)	3 (43)	0 (0)
Agree/strongly agree	18 (51)	2 (29)	5 (71)	5 (71)	2 (29)	4 (57)

Textbox 1. Themes, insights, and illustrative quotes describing the feasibility of the prediction tool.

Taking Action

Insights: The prediction tool led to no change in clinical action; the prediction tool can increase confidence in clinical action; the prediction tool can help prioritize clinical action and patients; and being confirmed supports decision-making.

- **Quote 1:** *Well, it hasn't changed my current decision, but the basis is much better, and I can easily see that it has helped me.* [Case 3, Electrophysiologist #7]
- **Quote 2:** *If you are in a busy situation where many transmissions have arrived and the technician and I have to maneuver and prioritize, there is no doubt that we will concentrate on those with high-risk predictions.* [Electrophysiologist #5]
- **Quote 3:** *This [tool prediction] is something that might make me react a little more aggressively. […] Now I've been told that he's actually more likely to get an episode within the next month than he's not getting an episode […] if our program is fully booked, both today and tomorrow, and the day after tomorrow, but on Friday we have a time. Then I kind of have to make a trade off if I really want to spare him a shock. Which may turn into a lot of shocks.* [Case 1, Electrophysiologist #2]

Decision-Making

Insights: The prediction tool predictions served as a second opinion; the prediction tool supported gathering of thoughts; the overall presentation of the prediction tool needs to be easily translatable to clinical relevance; and being confirmed supports decision-making.

- **Quote 4:** *So, I agree with the conclusion, it was also my feeling that I would be a little worried about this patient.* [Case 3, Electrophysiologist #6]
- **Quote 5:** *But then if it is you now have to convince some [other electrophysiologists] that they should ablate her, then instead of saying that I think so, you can argue that the algorithm thinks so too. So, in that way you can say that you can get an extra view of it.* [Case 3, Electrophysiologist #3]
- **Quote 6:** *In that way, the algorithm can be a support because it helps to gather thoughts about things that play a role in whether a person gets a new arrhythmia.* [Case 3, Electrophysiologist #5]
- **Quote 7:** *Yes, I think again that if you present 58.6% then it expresses an accuracy that you may have difficulty navigating with. I know it from other areas in the medical world, the thing about expressing something with a decimal number, it expresses an accuracy for which there may be no evidence at all […] I have a hard time relating to the number […] it's problematic to translate that into something clinically relevant.* [Case 1, Electrophysiologist #5]
- **Quote 8:** *I agree with what the alarm tells me, but I don't think it has helped me very much right here.* [Case 5, Electrophysiologist #5]

Visualization

Insights: The prediction tool should provide actionable parameters; showing parameters enables confirmation and agreement; showing parameters enables in-situ validation of algorithmic inputs and the prediction tool result; the prediction tool performs only as good as the data it bases its predictions on; transparency about the algorithmic data input helped raise confidence and trust; and showing important parameters is more important than showing the output probability.

- **Quote 9:** *To list what counts for and what counts against, makes really good sense. That's also how it works in my head.* [Case 3, Electrophysiologist #7]
- **Quote 10:** *I think it's super good, I actually think it's really pedagogical, I like it. Because, in reality this is how it confirms the result. It's basically the same empirical data that you have in your mind: You say "okay, is this a case where we have to do something?" It sums up some assumptions that you have made yourself, and in that way, I actually think you are confirmed more than if you have a green or red light.* [Electrophysiologist #4]
- **Quote 11:** *It's very nice to see that the algorithm reacts on the same parameters that I've discovered myself ... So it's nice to see that I agree with it. You could say that it's supporting and it's safe to know, that it also says there was something here.* [Case 3, Electrophysiologist #1]
- **Quote 12:** *What's happening here is that the ICD detects that the patient has VT, and then the prediction tool bases it's predictions on that. But it's not entirely correct, because the device has recently been re-programmed to sense everything.* [Case 1, Electrophysiologist #6]
- **Quote 13:** *This case has nothing to do with risk of VT/VF [...] it's the second thing I look at. No, here I won't use it [the prediction tool].* [Case 4, Electrophysiologist #1]

Time Saving

Insights: The prediction tool can speed up decision-making when trust is established; the prediction tool can reduce workload when trust is established; the prediction tool can reduce information search time when no or low risk is predicted; the prediction tool can substitute patient input; and showing important parameters enables work redelegation to technicians and nurses.

- **Quote 14:** *It will give me a much better basis for decision making and I actually think it will save me a lot of time. Just like with all other new technology based on machine learning: the first 2 months I sit and read through to see what I have, but in month 3, I will look at the output alone. Because then I trust that it has pulled out what is appropriate, and then it starts saving me all the work I did in the beginning. But for everyone, it is that there is a phase for you personally to find out if this brings you further. […] I really think I would have come to the decision faster if I had seen this first.* [Case 3, Electrophysiologist #7]

- **Quote 15:** *I might reach a decision faster with this system if I can't get a hold of the patient i.e., if the patient does not pick up the phone. Then it could well be that I look at the alarm and say "well, yes okay there is low risk."* [Case 2, Electrophysiologist #6]
- **Quote 16:** *You could make a scenario where the technician first looks at it [the prediction tool] and says ... okay there are those parameters and there is electrical storm, so we call in the patient and the doctor does not have to look at the transmission. That would support our workflow.* [Case 1, Electrophysiologist #4]

In Which Cases Does the Prediction Tool Support Clinical Decision-Making?

In 23 (66%) of the case walkthroughs, the electrophysiologists agreed that the prediction tool supported their decision-making, whereas in 8 (23%) of the walkthroughs they disagreed. Finding the prediction tool supporting was particularly pertinent in both patient cases 2 and 3, where 6 (86%) of the electrophysiologists agreed, and the prediction tool was found to assist decision-making by confirming the electrophysiologists' clinical evaluations and expectations of an increasing risk of VF/VT (Textbox 1, Quote 4). On the contrary, when the electrophysiologists were focused on predicting arrhythmias other than VT/VF, the prediction tool was deemed less useful, and answers were more heterogenous (Case 4 and Case 5). Some electrophysiologists said that the predictions served as a second opinion (Textbox 1, Quote 5) and that the prediction tool was helpful for collecting arguments that supported the electrophysiologists when trying to "gather thoughts" about potential VT/VF occurrences (Textbox 1, Quote 6). Nevertheless, some electrophysiologists found that showing the probability score as a percentage with decimals created uncertainty, and the naming of parameters was sometimes found difficult to interpret (Textbox 1, Quote 7).

Is Visualization of Important Parameters Useful?

The prediction tool's visualization of the most important parameters in the prediction of increased or decreased probability of VT/VF arrhythmia was found useful when the electrophysiologists agreed with the parameters presented. In patient cases 2 and 3, 6 (86%) and 5 (71%) of the electrophysiologists agreed that showing important parameters supported their decision-making. However, when the parameters represented poor data quality (Textbox 1, Quote 12), agreement was lower, for example, in patient Case 1 (43% agreed, 43% disagreed), or when the electrophysiologists were focused on predicting arrhythmias (Case 4 and Case 5) other than what the prediction tool was designed for (Textbox 1, Quote 13).

In general, presentation of important parameters provided explainability and supported decision-making by resembling the clinical interpretation process of what counts for or against the occurrence of VT/VF (Textbox 1, Quote 9). Several electrophysiologists found that visualization of important parameters created more confidence in the prediction tool than the probability score alone as the tool summed up many of the same assumptions that the electrophysiologists already had (Textbox 1, Quote 10). Listing the algorithm's important parameters also enabled electrophysiologists to do in-situ validations of the prediction tool's predictions by interpreting the data against the patient case (Textbox 1, Quote 11). However, in some cases, the electrophysiologists found that the parameters were based on wrong data from the ICD transmission. In those cases, it enabled electrophysiologists to check if the prediction tool based its predictions on wrong or poor data quality and to decide whether to trust the predictions or not (Textbox 1, Quote 12).

Does the Prediction Tool Reduce Time for Decision-Making?

The electrophysiologists found that the prediction tool could enable a reduction in time for decision-making in cases where they trusted the predictions. Moreover, 5 (71%) of the electrophysiologists agreed that the prediction tool can help reach a decision faster (Case 2 and Case 3). However, agreement was lower (29% in Case 1 and 57% in Case 5) when predictions were found to be uncertain or less useful for handling patients.

Several of the electrophysiologists expressed that once they become familiar with the system, they expect the AI tool will speed up decision-making and reduce the diagnostic workload. This indicates that establishing trust in AI predictions is essential. One of the electrophysiologists explained how time can be saved when personal trust in the prediction tool is developed (Textbox 1, Quote 14).

Across all cases, several electrophysiologists found that the probability score and the presentation of important parameters can reduce information search time. Typically, electrophysiologists must retrieve valuable information by clicking through multiple webpages in the ICD manufacturer's web-based system, which the prediction tool summarizes in a table. Some electrophysiologists also speculated that the tool could support decision-making when patient input is inaccessible, such as when a patient does not answer the phone (Textbox 1, Quote 15). Other electrophysiologists considered that the tool can support workflow and reduce unnecessary time consumption for electrophysiologists by delegating decision-making to the technician (Textbox 1, Quote 16).

Clinician Preimplementation Perspectives of the Prediction Tool: Acceptability, Adoption, Demand, and Implementation

Acceptability

Acceptability of the prediction tool was high when patient cases concerned VT/VF, as the risk predictions were found to be relevant. However, several electrophysiologists had expectations that the prediction tool would bring new and groundbreaking insights (Textbox 2, Quote 1) to support or challenge their decisions on which action to take. In cases where the task-technology fit was lower (Case 1, Case 4, and Case 5), acceptability was also lower (Textbox 2, Quote 2). For some of the electrophysiologists, the prediction tool was therefore considered "nice to have" rather than "need to have" (Textbox 2, Quote 3), while most of the electrophysiologists recognized the potential of the prediction tool. Some of the

electrophysiologists considered the tool useful for standardizing decision-making across the electrophysiologist team by avoiding individual influences from recent experiences and thus achieving harmonization of individual treatment (Textbox 2, Quote 4).

Adoption

There was consensus that high precision is important for prediction tool adoption to happen. Several of the electrophysiologists emphasized that the positive or negative predictive value should be as unambiguous as possible, showing either low or high risk when the alarm is raised (Textbox 2, Quote 5). Other electrophysiologists emphasized that false positives or negatives hinder adoption, which they explained to be the case for the adoption of OptiVol (an early warning alarm for fluid-related decompensation). Here, the electrophysiologist team decided not to use it due to too many false positives (Textbox 2, Quote 6). Several electrophysiologists explained that acceptance and clinical adoption are collectively decided based on team experiences from real-world use (Textbox 2, Quote 7) and from experiencing that the prediction tool actually confirms decisions in everyday clinical practice (Textbox 2, Quote 8). Adoption can also be achieved through building trust in the tool by means of validation studies. Participants explained that trust is a precondition for adoption, which can be achieved by documenting effects in a randomized clinical trial and through algorithm validation in peer-reviewed journals (Textbox 2, Quote 9).

Demand

Several of the electrophysiologists emphasized that there is a high demand for workflow support in remote monitoring of cardiac device patients. They found the prediction tool useful for supporting more efficient prioritization and identification of important patient cases (Textbox 2, Quote 10). Others described the demand for screening support among the increasing number of nonspecialized hospitals where fewer electrophysiologists are at work. For example, the prediction tool could support technicians doing the initial prioritization work more effectively and efficiently; the prediction tool could decrease electrophysiologists' patient information search time when handed over from technicians; and the prediction tool could function as "data help" by enabling junior doctors to get a form of senior help by consulting the tool (Textbox 2, Quote 11).

Implementation

To ensure successful implementation, some electrophysiologists described how remote monitoring clinics may want to be able to adjust the threshold of the prediction tool to fit local workflows and prioritization rules. For example, technicians and electrophysiologists should be able to configure the prediction tool and decide on related actions, such as "no need to take action" or "need to contact the patient." Relatedly, several electrophysiologists explained that indications of low-risk patients are especially useful in supporting clinicians in handling low-risk transmissions (Textbox 2, Quote 12). Moreover, the electrophysiologists explained that the intention of using the prediction tool is dependent on easy access, as well as how well it presents data and alleviates the need for clicking through several web pages in remote monitoring systems (Textbox 2, Quote 13). Some electrophysiologists added that it is practical that the algorithm uses data already available in remote monitoring systems, which are used daily for decision-making in the clinic. Knowing the data creates transparency and enables in-situ validation of the correctness of the probability score, thereby increasing the likelihood for success with implementation of the prediction tool (Textbox 2, Quote 14).

Textbox 2. Themes, insights, and illustrative quotes describing clinicians' preimplementation perspectives of the prediction tool.

Acceptability

Insights: Expectations that the prediction tool would bring new and more groundbreaking insights; overall usefulness of the prediction tool is "nice to have;" clear purpose is decisive for acceptability; intension of use is tied to the prediction tool's task-technology fit; harmonizing individual treatment; and avoiding being influenced by recent experiences and reducing individual bias.

- **Quote 1:** *[…] it confirms the assessment you make, and that's fine, but it's not something groundbreaking, and that's okay too.* [Interview, Electrophysiologist #5]
- **Quote 2:** *[…] I'm a little disappointed with the alarm, because for the cases I have looked at, the alarm has not given me much. […] if you had some cases with some 'meat' on such as a couple of treatment requiring VTs, I actually do think that getting a number, a risk score, will enable to better estimating the problem. I think that can be valuable […] I just think the cases were wrong. If you want to show that this algorithm gives value and then bring 2 cases with AF problems, which the algorithm does not handle, then that's not optimal.* [Interview, Electrophysiologist #2]
- **Quote 3:** *It's always nice when something is supportive, I would say, but isn't it a "nice to have" and not a "need to have"?* [Interview, Electrophysiologist #4]
- **Quote 4:** *20 years of experience or not. Perhaps, the advantage of the algorithm is that it is not influenced by what the individual clinician has experienced within the last month, and in this way helping to make more uniform conclusions.* [Interview, Electrophysiologist #5]

Adoption

Insights: High precision is important; false positives hinder adoption; using the prediction tool and getting confirmation in real life creates trust and enables adoption; the clinical team needs to decide on use; a randomized clinical trial is a precondition for acceptability; and algorithm validation supports trust.

- **Quote 5:** *If you want to come out with this, it must be something with a positive predictive value that is really good, so that you don't get a lot of nonsense that you can't use. The alarm should only be raised when there really is something.* [Interview, Electrophysiologist #3]
- **Quote 6:** *It needs to be easily accessible and we [team of electrophysiologists] have to agree that we trust it [the prediction tool]. We just have to say that yes it looks right. For example, the Optivol alarm had too many false positives, which gave a lot of extra work and everything, and we actually chose not to use it because there were too many sources of error, and you only really discover that when you work with it [new algorithms].* [Interview, Electrophysiologist #1]
- **Quote 7:** *I would say that it [prediction tool] would be an instrument that would have to be accepted in our group and then you would find it valuable when we all agree to take the red alarms first, and in that way use it to prioritize a bit.* [Interview, Electrophysiologist #5]
- **Quote 8:** *I just think I should see that it confirms our decisions in enough cases - then I would feel comfortable about colleagues leaning on it […] There is something about trying it out, you know how it is.* [Interview, Electrophysiologist #4]
- **Quote 9:** *Published studies of the algorithm would increase confidence yes, because then you know that someone with an understanding of making these models have said that it looks okay; someone externally who have validated it.* [Electrophysiologist #1]

Demand

Insights: Supporting better workflows; demand for prioritization and identification of important patient cases; increasing demand for screening tools in nonspecialist hospitals; demand for decreasing electrophysiologist's information search time; supporting nurses and technicians to do prioritization work; and "data help" that enables junior doctors to get senior help.

- **Quote 10:** *When transmissions come in, it's almost an unsorted list of transmissions […] The list is unprocessed, so with the algorithm it takes it a step further by nuancing what comes into CareLink [Medtronic's remote monitoring dashboard] with some semi-quantitative markings. And, if it is reliable, then it would be valuable. Partly because you don't overlook anything, and partly because you are confirmed that we must take these patients first, because we have experience that there can be trouble here.* [Interview, Electrophysiologist #5]
- **Quote 11:** *You could say that in this way, the young doctor can do without getting senior help by actually getting data help.* [Interview, Electrophysiologist #2]

Implementation

Insights: Demand for adjusting the threshold to local prioritization rules; clinics need to be able to configure the cutoff and threshold; making the prediction tool easily accessible and integrated in the list of transmissions supports the workflow; intention of use is dependent on easy access; and it is practical that the algorithm uses data that are already familiar to the clinicians.

- **Quote 12:** *Electrophysiologists don't bother to hear about it if it is below a certain percentage […]. We have adopted some rules, e.g. if you have a patient and she has got a shock, and gets rare therapies and it goes over, then we don't need to hear about it because we think there is not a big risk. You might well imagine that introducing this alarm will support handling low-risk transmissions.* [Interview, Electrophysiologist #3]
- **Quote 13:** *If it's easily presented and you don't have to go in and look through 4 pages and such and if it was on the front page and brought up "number of episodes" and information like that - if you could easily retrieve the information [from the prediction tool] or if it was printed on the list of transmissions we are working on, then it would also be a great help.* [Interview, Electrophysiologist #1]

- **Quote 14:** *What one would emphasize, is that the algorithm uses the same data that the clinician uses i.e. it's the same data, just integrated according to a formula that clinicians do not currently have available.* [Interview, Electrophysiologist #5]

Discussion

Tackling the Sociotechnical Challenges of ML-Based Tools in Health Care

In bridging the sociotechnical gap between the development of ML-based tools and clinical implementation, this study explored the feasibility and clinical perspectives of using a prediction tool for improved workflows in ICD remote monitoring. We found that the feasibility of the ML-based tool is promising when the intended use of the tool is aligned with expectations, that is, by providing support for decision-making, visualizing useful information, and reducing time spent. The results also show that an actionable prediction tool is one that presents the reason for why the algorithm deemed as it did, such as in this study, by highlighting important data to be used for clinical evaluation and enabling clinicians to assess the algorithm's outcome against their own evaluation [31,33].

However, the current prediction tool did not lead to change in clinical action, suggesting that ML and explainability techniques do not outperform specialized and experienced electrophysiologist evaluations, but at best confirm and support the interpretation of complex ICD device information along with a promise for a less time-consuming clinical workflow.

The contribution of this paper lies in the implications of the qualitative results suggesting that clinical end-users, clinical contexts, and workflows must be included throughout an overall iterative approach to design, development, and implementation. In the following sections, we will discuss the qualitative results concerning the sociotechnical challenges and implementation of ML-based tools for clinical decision support.

Expectations Need to Align With the Intended Use of AI

In cases where misalignment emerged between the electrophysiologists' expectations and intended use, the prediction tool was considered less useful and at best "nice to have" for clinical decision-making. For example, in cases where the ICD transmissions revolved around other types of arrhythmias than what the prediction tool was designed for and in cases where the electrophysiologists expected that the prediction tool should be capable of outperforming their own evaluation, disappointment was raised about the performance of the underlying AI algorithm. This aligns with recent studies that reported on physicians' high expectations and attitudes toward medical AI [22,23,51,65]. The challenge of managing expectations has been addressed by a growing number of studies aimed at providing an explanation of algorithmic decisions at the time of inference [36] and by developing user interfaces with expectation adjustment techniques [66]. Recently, researchers focused on the early human-AI onboarding process of pathologists and found that presenting a global view of a prediction tool and its capabilities, limitations, and biases is key to the formation of initial impressions and appropriate mental models [67]. This suggests that the development of so-called explainability in the user interface is important, but communicating the intended use of the prediction tool is imperative for acceptance in the clinic. To achieve alignment of expectations, training programs for clinicians are critical when implementing medical AI tools.

Trust Emerges From Real-World Use

Trust is another key factor for user acceptance and adoption of AI technologies. Trust is typically considered an issue in creating transparent and understandable algorithmic behavior, as opposed to seeing the prediction tool as a black box [55,68,69]. Extensive research on explainable AI and various approaches to achieve transparency have been suggested [11], yet experimental studies on whether these approaches achieve their intended effects in the real world are only just starting to emerge [38,39,69,70]. In this study, the electrophysiologists requested large-scale algorithm validation and prospective evaluations from clinical trials. However, an important observation was that trust in the prediction tool may only emerge from continuous use of the tool and from experiencing confirmation on individual evaluations in collaboration with the tool. There was general agreement among the electrophysiologists that visualization of the most important predictive parameters helped raise confidence and trust over time, and that adoption of the prediction tool would hinge on the collective decision among the team of electrophysiologists. Recent experimental studies have reported similar findings [69,71] and have demonstrated that adding an AI prediction tool to the clinical evaluation can increase clinician confidence [24]. The implication of understanding trust as emerging from real-world use is that when deploying medical AI in clinical settings, trust needs to be built bottom-up through weeks or months of trialing the new tool for clinicians to experience convincing reassurance. Therefore, initial implementation processes may benefit from simultaneous calibration and adaption of the tool to establish a human-AI partnership, and allowing the local team of clinicians to decide collectively how they choose to trust and use the tool in the clinic.

Accuracy is Dependent on Workflow and Context

While AI algorithms have been validated and have been shown to have similar or higher accuracy than humans, recent studies of AI deployment in clinical settings report that professional autonomy, workflow, and local sociotechnical factors have impacts on how accuracy is perceived and used in clinical practice [24,43,45-47,50-54]. Bruun et al [24] found that overall performance was positively impacted among clinicians using an AI-prediction tool for assessing progression in early stage dementia and that clinicians' professional autonomy impacts the use of medical AI in situated clinical practice. Additionally, the study by Beede et al [29] of a ML-based (deep learning) system used in clinics for the detection of diabetic eye disease indicated that several socioenvironmental factors, such as busy screening procedures, poor lighting conditions, and consideration of patient burden, have impacts on how AI accuracy is perceived in clinical screening practices. Similarly,

we found that high accuracy becomes relative to the electrophysiologist's evaluation of available information, the local circumstances, and the consequences that AI predictions have for taking action. For example, several electrophysiologists argued that AI prediction needs to be considered against patient-reported symptoms and that a full patient schedule may affect how the AI prediction is acted upon in practice. Moreover, in several cases, the electrophysiologists found the visualization of important parameters more useful than the prediction score itself. This indicates that AI accuracy needs to be understood as relational and dependent on available information and local workflows, which supports the vision of establishing a human-AI symbiosis that combines the predictive abilities of both the clinician and the AI prediction algorithm [32,33]. Finally, the wish for better visualization of data parameters over prediction accuracy indicates that the development of medical AI assistants needs to be carried out as close as possible to implementation in clinical practice with clinical end-users through iterative approaches [37,42,72] that can bridge the "AI chasm" [41] of scientifically sound algorithms and their use in meaningful real-world clinical applications.

Limitations

The findings in this study are limited to the small number of study participants and patient cases. One electrophysiologist (PKJ) participated in co-design workshops, resulting in potential positive bias. Patient cases were selected to represent diversity in prediction capabilities, rather than the distribution in clinical practice, which may weaken the generalizability of the results. Only cases where the prediction tool provided true-positive and true-negative prediction outcomes were used, which means that the clinical feasibility of ML in cases with false-positive and false-negative outcomes [73] have not been explored. Future studies are needed to assess the implications of false prediction outcomes, as well as conduct algorithmic validation similar to recent related studies [14-21]. Limitations also involve data availability, that is, the data set used may entail algorithmic bias [13] and the study participants may have been more positive toward innovative AI technology since all of the study participants were from a tertiary university hospital and constituted a rather homogenous group of highly specialized physicians. The AI studied has limitations, because only the random forest ML-based algorithm was evaluated with electrophysiologists. These types of methods are commonly used in medical applications [21,74,75] because of their high classification accuracy and capabilities for handling data with imbalanced classes [50] while providing easily accessible, if limited, global intelligibility through the visualization and ranking of parameter importance [55]. This work will benefit from being validated in a large-scale multicenter study with higher diversity in participating electrophysiologists and workflows. It will be imperative to conduct prospective clinical trials evaluating the algorithm against standard care with regard to workload, cost-effectiveness, and hard clinical endpoints.

Conclusions

This study shows that a tool based on ML for the prediction of VT/VF in remote monitoring of ICD patients has the potential to support electrophysiologists' decision-making. While the prediction tool was regarded as "nice to have" rather than "need to have" in its current form, the tool demonstrated potential for supporting clinical decision-making, as it provided reassurance, increased confidence, and indicated the potential for reducing information search time, as well as enabled delegation of decisions to nurses and technicians. The findings also indicate that trust in the prediction tool, acceptable data quality, and clearly defined intended use are decisive for end-user acceptance and that adoption hinges on successful clinical implementation. This suggests that clinical end-users' sociotechnical contexts and workflows need to be taken into consideration early on and continuously throughout a participatory design process to address the sociotechnical gap between the development and implementation of medical AI in cardiac care.

Acknowledgments

We wish to thank the participating electrophysiologists at the Department of Cardiology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. The research and technology developed (working title: SafeHeart) were supported by the European Data Pitch Innovation Program H2020–732506 and led by TOA.

Conflicts of Interest

TOA is a co-founder of Vital Beats, which has commercial interests in the technology under investigation. MKHH and CV were full-time employees of Vital Beats. JHS, SZD, MCHL, and SM are affiliated with Vital Beats as advisors or independent researchers. The authors have no other conflicts of interest to disclose.

Multimedia Appendix 1
Questionnaire used before the electrophysiologists are presented with the prediction tool results.
[PDF File (Adobe PDF File), 79 KB - humanfactors_v8i4e26964_app1.pdf]

Multimedia Appendix 2
Questionnaire after the electrophysiologists have received the prediction tool results.
[PDF File (Adobe PDF File), 103 KB - humanfactors_v8i4e26964_app2.pdf]

Multimedia Appendix 3
The semistructured interview guide.
[PDF File (Adobe PDF File), 53 KB - humanfactors_v8i4e26964_app3.pdf]

References

1. Al-Khatib SM, Stevenson WG, Ackerman MJ, Bryant WJ, Callans DJ, Curtis AB, et al. 2017 AHA/ACC/HRS Guideline for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol 2018 Oct 02;72(14):e91-e220 [FREE Full text] [doi: 10.1016/j.jacc.2017.10.054] [Medline: 29097296]
2. Nanthakumar K, Epstein AE, Kay GN, Plumb VJ, Lee DS. Prophylactic implantable cardioverter-defibrillator therapy in patients with left ventricular systolic dysfunction: a pooled analysis of 10 primary prevention trials. J Am Coll Cardiol 2004 Dec 07;44(11):2166-2172 [FREE Full text] [doi: 10.1016/j.jacc.2004.08.054] [Medline: 15582314]
3. Varma N, Ricci RP. Telemedicine and cardiac implants: what is the benefit? Eur Heart J 2013 Jul 04;34(25):1885-1895 [FREE Full text] [doi: 10.1093/eurheartj/ehs388] [Medline: 23211231]
4. Mabo P, Victor F, Bazin P, Ahres S, Babuty D, Da Costa A, COMPAS Trial Investigators. A randomized trial of long-term remote monitoring of pacemaker recipients (the COMPAS trial). Eur Heart J 2012 May;33(9):1105-1111 [FREE Full text] [doi: 10.1093/eurheartj/ehr419] [Medline: 22127418]
5. Landolina M, Perego GB, Lunati M, Curnis A, Guenzati G, Vicentini A, et al. Remote monitoring reduces healthcare use and improves quality of care in heart failure patients with implantable defibrillators: the evolution of management strategies of heart failure patients with implantable defibrillators (EVOLVO) study. Circulation 2012 Jun 19;125(24):2985-2992 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.088971] [Medline: 22626743]
6. Saxon LA, Hayes DL, Gilliam FR, Heidenreich PA, Day J, Seth M, et al. Long-term outcome after ICD and CRT implantation and influence of remote device follow-up: the ALTITUDE survival study. Circulation 2010 Dec 07;122(23):2359-2367 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.960633] [Medline: 21098452]
7. Crossley GH, Boyle A, Vitense H, Chang Y, Mead RH, CONNECT Investigators. The CONNECT (Clinical Evaluation of Remote Notification to Reduce Time to Clinical Decision) trial: the value of wireless remote monitoring with automatic clinician alerts. J Am Coll Cardiol 2011 Mar 08;57(10):1181-1189 [FREE Full text] [doi: 10.1016/j.jacc.2010.12.012] [Medline: 21255955]
8. Cronin EM, Ching EA, Varma N, Martin DO, Wilkoff BL, Lindsay BD. Remote monitoring of cardiovascular devices: a time and activity analysis. Heart Rhythm 2012 Dec;9(12):1947-1951 [FREE Full text] [doi: 10.1016/j.hrthm.2012.08.002] [Medline: 22864266]
9. Ricci RP, Morichelli L, D'Onofrio A, Calò L, Vaccari D, Zanotto G, et al. Manpower and outpatient clinic workload for remote monitoring of patients with cardiac implantable electronic devices: data from the HomeGuide Registry. J Cardiovasc Electrophysiol 2014 Nov 28;25(11):1216-1223 [FREE Full text] [doi: 10.1111/jce.12482] [Medline: 24964380]
10. Andersen TO, Nielsen KD, Moll J, Svendsen JH. Unpacking telemonitoring work: Workload and telephone calls to patients in implanted cardiac device care. Int J Med Inform 2019 Sep;129:381-387 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.06.021] [Medline: 31445281]
11. Gustafson, PE. Implementing Remote Monitoring of Cardiac Implantable Electronic Devices: The Clinical Experience from One Center in Sweden. The Journal of Innovations in Cardiac Rhythm Management 2014;5:1733-1739 [FREE Full text] [doi: 10.19102/icrm.2014.050808]
12. Patel B, Sengupta P. Machine learning for predicting cardiac events: what does the future hold? Expert Rev Cardiovasc Ther 2020 Feb 23;18(2):77-84 [FREE Full text] [doi: 10.1080/14779072.2020.1732208] [Medline: 32066289]
13. Crowley RJ, Tan YJ, Ioannidis JPA. Empirical assessment of bias in machine learning diagnostic test accuracy studies. J Am Med Inform Assoc 2020 Jul 01;27(7):1092-1101 [FREE Full text] [doi: 10.1093/jamia/ocaa075] [Medline: 32548642]
14. Lagu T, Pekow PS, Shieh M, Stefan M, Pack QR, Kashef MA, et al. Validation and Comparison of Seven Mortality Prediction Models for Hospitalized Patients With Acute Decompensated Heart Failure. Circ: Heart Failure 2016 Aug;9(8):e002912 [FREE Full text] [doi: 10.1161/circheartfailure.115.002912]
15. Pan Y, Fu M, Cheng B, Tao X, Guo J. Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction on the Internet of Medical Things Platform. IEEE Access 2020;8:189503-189512 [FREE Full text] [doi: 10.1109/access.2020.3026214]
16. Eltrass AS, Tayel MB, Ammar AI. A new automated CNN deep learning approach for identification of ECG congestive heart failure and arrhythmia using constant-Q non-stationary Gabor transform. Biomedical Signal Processing and Control 2021 Mar;65:102326 [FREE Full text] [doi: 10.1016/j.bspc.2020.102326]
17. Hammad M, Iliyasu AM, Subasi A, Ho ESL, El-Latif AAA. A Multitier Deep Learning Model for Arrhythmia Detection. IEEE Trans. Instrum. Meas 2021;70:1-9 [FREE Full text] [doi: 10.1109/tim.2020.3033072]
18. Shouval R, Hadanny A, Shlomo N, Iakobishvili Z, Unger R, Zahger D, et al. Machine learning for prediction of 30-day mortality after ST elevation myocardial infraction: An Acute Coronary Syndrome Israeli Survey data mining study. Int J Cardiol 2017 Nov 01;246:7-13 [FREE Full text] [doi: 10.1016/j.ijcard.2017.05.067] [Medline: 28867023]

19. Karnan H, Natarajan S, Manivel R. Human machine interfacing technique for diagnosis of ventricular arrhythmia using supervisory machine learning algorithms. Concurrency Computat Pract Exper 2018 Oct 19;33(4):e5001 [FREE Full text] [doi: 10.1002/cpe.5001]
20. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. The Lancet 2019 Sep 07;394(10201):861-867 [FREE Full text] [doi: 10.1016/S0140-6736(19)31721-0] [Medline: 31378392]
21. Shakibfar S, Krause O, Lund-Andersen C, Aranda A, Moll J, Andersen TO, et al. Predicting electrical storms by remote monitoring of implantable cardioverter-defibrillator patients using machine learning. Europace 2019 Feb 01;21(2):268-274 [FREE Full text] [doi: 10.1093/europace/euy257] [Medline: 30508072]
22. Oh S, Kim JH, Choi S, Lee HJ, Hong J, Kwon SH. Physician Confidence in Artificial Intelligence: An Online Mobile Survey. J Med Internet Res 2019 Mar 25;21(3):e12422 [FREE Full text] [doi: 10.2196/12422] [Medline: 30907742]
23. Maassen O, Fritsch S, Palm J, Deffge S, Kunze J, Marx G, et al. Future Medical Artificial Intelligence Application Requirements and Expectations of Physicians in German University Hospitals: Web-Based Survey. J Med Internet Res 2021 Mar 05;23(3):e26646 [FREE Full text] [doi: 10.2196/26646] [Medline: 33666563]
24. Bruun M, Frederiksen KS, Rhodius-Meester HFM, Baroni M, Gjerum L, Koikkalainen J, et al. Impact of a clinical decision support tool on prediction of progression in early-stage dementia: a prospective validation study. Alzheimers Res Ther 2019 Mar 20;11(1):25 [FREE Full text] [doi: 10.1186/s13195-019-0482-3] [Medline: 30894218]
25. Manz C, Chivers C, Liu M, Regli SB, Changolkar S, Evans CN, et al. Prospective validation of a machine learning algorithm to predict short-term mortality among outpatients with cancer. JCO 2020 May 20;38(15_suppl):2009-2009 [FREE Full text] [doi: 10.1200/jco.2020.38.15_suppl.2009]
26. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med 2018;1:39 [FREE Full text] [doi: 10.1038/s41746-018-0040-6] [Medline: 31304320]
27. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA 2016 Dec 13;316(22):2402-2410 [FREE Full text] [doi: 10.1001/jama.2016.17216] [Medline: 27898976]
28. Kang J, Morin O, Hong JC. Closing the Gap Between Machine Learning and Clinical Cancer Care-First Steps Into a Larger World. JAMA Oncol 2020 Nov 01;6(11):1731-1732 [FREE Full text] [doi: 10.1001/jamaoncol.2020.4314] [Medline: 32970129]
29. Beede E, Baylor E, Hersch F, Iurchenko A, Wilcox L, Ruamviboonsuk P, et al. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. In: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: 2020 CHI Conference on Human Factors in Computing Systems; April 25-30, 2020; Honolulu, HI p. 1-12. [doi: 10.1145/3313831.3376718]
30. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. N Engl J Med 2017 Jun 29;376(26):2507-2509 [FREE Full text] [doi: 10.1056/NEJMp1702071] [Medline: 28657867]
31. Lindsell CJ, Stead WW, Johnson KB. Action-Informed Artificial Intelligence-Matching the Algorithm to the Problem. JAMA 2020 Jun 02;323(21):2141-2142. [doi: 10.1001/jama.2020.5035] [Medline: 32356878]
32. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56 [FREE Full text] [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
33. Lee J. Is Artificial Intelligence Better Than Human Clinicians in Predicting Patient Outcomes? J Med Internet Res 2020 Aug 26;22(8):e19918 [FREE Full text] [doi: 10.2196/19918] [Medline: 32845249]
34. Cabitza F, Campagner A, Balsano C. Bridging the "last mile" gap between AI implementation and operation: "data awareness" that matters. Ann Transl Med 2020 Apr;8(7):501 [FREE Full text] [doi: 10.21037/atm.2020.03.63] [Medline: 32395545]
35. Coiera E. The Last Mile: Where Artificial Intelligence Meets Reality. J Med Internet Res 2019 Nov 08;21(11):e16323 [FREE Full text] [doi: 10.2196/16323] [Medline: 31702559]
36. Abdul A, Vermeulen J, Wang D, Lim BY, Kankanhalli M. Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda. In: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, QC, Canada p. 1-18. [doi: 10.1145/3173574.3174156]
37. Wolf CT. Explainability scenarios: towards scenario-based XAI design. In: IUI '19: Proceedings of the 24th International Conference on Intelligent User Interfaces. 2019 Presented at: 24th International Conference on Intelligent User Interfaces; March 17-20, 2019; Marina del Ray, CA p. 252-257. [doi: 10.1145/3301275.3302317]
38. Miller T. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence 2019 Feb;267:1-38 [FREE Full text] [doi: 10.1016/j.artint.2018.07.007]
39. Poursabzi-Sangdeh F, Goldstein DG, Hofman JM, Vaughan JW, Wallach H. Manipulating and Measuring Model Interpretability. In: CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama Japan p. 1-52. [doi: 10.1145/3411764.3445315]

40. Coiera EW. Artificial intelligence in medicine: the challenges ahead. J Am Med Inform Assoc 1996;3(6):363-366 [FREE Full text] [doi: 10.1136/jamia.1996.97084510] [Medline: 8930853]
41. Keane PA, Topol EJ. With an eye to AI and autonomous diagnosis. NPJ Digit Med 2018;1:40 [FREE Full text] [doi: 10.1038/s41746-018-0048-y] [Medline: 31304321]
42. Quiroz JC, Laranjo L, Kocaballi AB, Berkovsky S, Rezazadegan D, Coiera E. Challenges of developing a digital scribe to reduce clinical documentation burden. NPJ Digit Med 2019 Nov 22;2(1):114 [FREE Full text] [doi: 10.1038/s41746-019-0190-1] [Medline: 31799422]
43. Kocaballi AB, Ijaz K, Laranjo L, Quiroz JC, Rezazadegan D, Tong HL, et al. Envisioning an artificial intelligence documentation assistant for future primary care consultations: A co-design study with general practitioners. J Am Med Inform Assoc 2020 Nov 01;27(11):1695-1704 [FREE Full text] [doi: 10.1093/jamia/ocaa131] [Medline: 32845984]
44. Petkovic D, Altman R, Wong M, Vigil A. Improving the explainability of Random Forest classifier - user centered approach. Pac Symp Biocomput 2018;23:204-215 [FREE Full text] [Medline: 29218882]
45. Benda NC, Das LT, Abramson EL, Blackburn K, Thoman A, Kaushal R, et al. "How did you get to this number?" Stakeholder needs for implementing predictive analytics: a pre-implementation qualitative study. J Am Med Inform Assoc 2020 May 01;27(5):709-716 [FREE Full text] [doi: 10.1093/jamia/ocaa021] [Medline: 32159774]
46. Sendak M, Elish MC, Gao M, Futoma J, Ratliff W, Nichols M, et al. "The human body is a black box": supporting clinical decision-making with deep learning. In: FAT '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. 2020 Presented at: 2020 Conference on Fairness, Accountability, and Transparency; January 27-30, 2020; Barcelona, Spain p. 99-109. [doi: 10.1145/3351095.3372827]
47. Elish MC, Watkins EA. Repairing Innovation: A Study of Integrating AI in Clinical Care. Data Society. 2020. URL: https://datasociety.net/wp-content/uploads/2020/09/Repairing-Innovation-DataSociety-20200930-1.pdf [accessed 2021-11-18]
48. Andersen TO, Nunes F, Wilcox L, Kaziunas E, Matthiesen S, Magrabi F. Realizing AI in Healthcare: Challenges Appearing in the Wild. In: CHI EA '21: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan p. 1-5. [doi: 10.1145/3411763.3441347]
49. Ginestra JC, Giannini HM, Schweickert WD, Meadows L, Lynch MJ, Pavan K, et al. Clinician Perception of a Machine Learning–Based Early Warning System Designed to Predict Severe Sepsis and Septic Shock*. Critical Care Medicine 2019;47(11):1477-1484 [FREE Full text] [doi: 10.1097/ccm.0000000000003803]
50. Barda AJ, Horvat CM, Hochheiser H. A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. BMC Med Inform Decis Mak 2020 Oct 08;20(1):257 [FREE Full text] [doi: 10.1186/s12911-020-01276-x] [Medline: 33032582]
51. Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. A lesson in implementation: A pre-post study of providers' experience with artificial intelligence-based clinical decision support. Int J Med Inform 2020 May;137:104072 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.104072] [Medline: 32200295]
52. Sandhu S, Lin AL, Brajer N, Sperling J, Ratliff W, Bedoya AD, et al. Integrating a Machine Learning System Into Clinical Workflows: Qualitative Study. J Med Internet Res 2020 Nov 19;22(11):e22421 [FREE Full text] [doi: 10.2196/22421] [Medline: 33211015]
53. Wang D, Wang L, Zhang Z, Wang D, Zhu H, Gao Y, et al. “Brilliant AI Doctor” in Rural Clinics: Challenges in AI-Powered Clinical Decision Support System Deployment. In: CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: 2021 CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan p. 1-18. [doi: 10.1145/3411764.3445432]
54. Yang Q, Steinfeld A, Zimmerman J. Unremarkable AI: Fitting Intelligent Decision Support into Critical, Clinical Decision-Making Processes. In: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland p. 1-11. [doi: 10.1145/3290605.3300468]
55. Ribeiro MT, Singh S, Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In: KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016 Presented at: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA p. 1135-1144. [doi: 10.1145/2939672.2939778]
56. Andersen T, Bjørn P, Kensing F, Moll J. Designing for collaborative interpretation in telemonitoring: re-introducing patients as diagnostic agents. Int J Med Inform 2011 Aug;80(8):e112-e126 [FREE Full text] [doi: 10.1016/j.ijmedinf.2010.09.010] [Medline: 21067968]
57. Altman NS. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. The American Statistician 1992 Aug;46(3):175 [FREE Full text] [doi: 10.2307/2685209]
58. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann. Statist 2001 Oct 1;29(5):1189-1232 [FREE Full text] [doi: 10.1214/aos/1013203451]
59. Freund Y, Schapire R. A short introduction to boosting. Journal of Japanese Society for Artificial Intelligence 1999 Sep;14(5):771-780 [FREE Full text]

60. Chang CC, Lin CJ. LIBSVM. ACM Trans. Intell. Syst. Technol 2011 Apr;2(3):1-27 [FREE Full text] [doi: 10.1145/1961189.1961199]
61. Karim F, Majumdar S, Darabi H, Chen S. LSTM Fully Convolutional Networks for Time Series Classification. IEEE Access 2018;6:1662-1669 [FREE Full text] [doi: 10.1109/access.2017.2779939]
62. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
63. Li AC, Kannry JL, Kushniruk A, Chrimes D, McGinn TG, Edonyabo D, et al. Integrating usability testing and think-aloud protocol analysis with "near-live" clinical simulations in evaluating clinical decision support. Int J Med Inform 2012 Nov;81(11):761-772. [doi: 10.1016/j.ijmedinf.2012.02.009] [Medline: 22456088]
64. Charmaz K. Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. London: Sage Publications; 2006.
65. Abdullah R, Fakieh B. Health Care Employees' Perceptions of the Use of Artificial Intelligence Applications: Survey Study. J Med Internet Res 2020 May 14;22(5):e17620 [FREE Full text] [doi: 10.2196/17620] [Medline: 32406857]
66. Kocielnik R, Amershi S, Bennett PN. Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-user Expectations of AI Systems. In: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland p. 1-14. [doi: 10.1145/3290605.3300641]
67. Cai CJ, Winter S, Steiner D, Wilcox L, Terry M. "Hello AI": Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. Proc. ACM Hum.-Comput. Interact 2019 Nov 07;3(CSCW):1-24. [doi: 10.1145/3359206]
68. LaRosa E, Danks D. Impacts on Trust of Healthcare AI. In: AIES '18: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. 2018 Presented at: 2018 AAAI/ACM Conference on AI, Ethics, and Society; February 2-3, 2018; New Orleans, LA p. 210-215. [doi: 10.1145/3278721.3278771]
69. Yin M, Wortman Vaughan J, Wallach H. Understanding the Effect of Accuracy on Trust in Machine Learning Models. In: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland p. 1-12. [doi: 10.1145/3290605.3300509]
70. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak 2020 Nov 30;20(1):310 [FREE Full text] [doi: 10.1186/s12911-020-01332-6] [Medline: 33256715]
71. Maniatopoulos G, Procter R, Llewellyn S, Harvey G, Boyd A. Moving beyond local practice: reconfiguring the adoption of a breast cancer diagnostic technology. Soc Sci Med 2015 Apr;131:98-106 [FREE Full text] [doi: 10.1016/j.socscimed.2015.02.036] [Medline: 25769108]
72. Wang D, Yang Q, Abdul A, Lim BY. Designing Theory-Driven User-Centric Explainable AI. In: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland p. 1-15. [doi: 10.1145/3290605.3300831]
73. Gajowniczek K, Grzegorczyk I, Ząbkowski T. Reducing False Arrhythmia Alarms Using Different Methods of Probability and Class Assignment in Random Forest Learning Methods. Sensors (Basel) 2019 Apr 02;19(7):1588 [FREE Full text] [doi: 10.3390/s19071588] [Medline: 30986930]
74. Azar AT, Elshazly HI, Hassanien AE, Elkorany AM. A random forest classifier for lymph diseases. Comput Methods Programs Biomed 2014 Feb;113(2):465-473 [FREE Full text] [doi: 10.1016/j.cmpb.2013.11.004] [Medline: 24290902]
75. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform Decis Mak 2011 Jul 29;11(1):51-13 [FREE Full text] [doi: 10.1186/1472-6947-11-51] [Medline: 21801360]

Abbreviations

AI: artificial intelligence
ICD: implantable cardioverter defibrillator
LSTM: long short-term memory
ML: machine learning
VF: ventricular fibrillation
VT: ventricular tachycardia

Edited by A Kushniruk; submitted 05.01.21; peer-reviewed by T Canares, V Montmirail; comments to author 02.03.21; revised version received 23.03.21; accepted 11.10.21; published 26.11.21.

Please cite as:
Matthiesen S, Diederichsen SZ, Hansen MKH, Villumsen C, Lassen MCH, Jacobsen PK, Risum N, Winkel BG, Philbert BT, Svendsen JH, Andersen TO
Clinician Preimplementation Perspectives of a Decision-Support Tool for the Prediction of Cardiac Arrhythmia Based on Machine Learning: Near-Live Feasibility and Qualitative Study
JMIR Hum Factors 2021;8(4):e26964
URL: https://humanfactors.jmir.org/2021/4/e26964
doi:10.2196/26964
PMID:34842528

Original Paper

Pediatric Clinicians' Use of Telemedicine: Qualitative Interview Study

Julia B Finkelstein[1*], MD, MPH; Elise S Tremblay[2*], MD, MPH; Melissa Van Cain[3], MD, MBI; Aaron Farber-Chen[4], MSN, RN, FNP-BC; Caitlin Schumann[4], MPH; Christina Brown[4], MSc; Ankoor S Shah[4,5], MD, PhD; Erinn T Rhodes[2], MD, MPH

[1]Department of Urology, Boston Children's Hospital, Boston, MA, United States

[2]Division of Endocrinology, Boston Children's Hospital, Boston, MA, United States

[3]Department of Medical Informatics, School of Community Medicine, University of Oklahoma, Tulsa, OK, United States

[4]Innovation and Digital Health Accelerator, Boston Children's Hospital, Boston, MA, United States

[5]Department of Ophthalmology, Boston Children's Hospital, Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Julia B Finkelstein, MD, MPH
Department of Urology
Boston Children's Hospital
300 Longwood Avenue
Boston, MA, 02115
United States
Phone: 1 6173557796
Fax: 1 6177300474
Email: julia.finkelstein@childrens.harvard.edu

Abstract

Background: Bedside manner describes how clinicians relate to patients in person. Telemedicine allows clinicians to connect virtually with patients using digital tools. Effective virtual communication or *webside manner* may require modifications to traditional bedside manner.

Objective: This study aims to understand the experiences of telemedicine providers with patient-to-provider virtual visits and communication with families at a single large-volume children's hospital to inform program development and training for future clinicians.

Methods: A total of 2 focus groups of pediatric clinicians (N=11) performing virtual visits before the COVID-19 pandemic, with a range of experiences and specialties, were engaged to discuss experiential, implementation, and practice-related issues. Focus groups were facilitated using a semistructured guide covering general experience, preparedness, rapport strategies, and suggestions. Sessions were digitally recorded, and the corresponding transcripts were reviewed for data analysis. The transcripts were coded based on the identified main themes and subthemes. On the basis of a higher-level analysis of these codes, the study authors generated a final set of key themes to describe the collected data.

Results: Theme consistency was identified across diverse participants, although individual clinician experiences were influenced by their specialties and practices. A total of 3 key themes emerged regarding the development of best practices, barriers to scalability, and establishing patient rapport. Issues and concerns related to privacy were salient across all themes. Clinicians felt that telemedicine required new skills for patient interaction, and not all were comfortable with their training.

Conclusions: Telemedicine provides benefits as well as challenges to health care delivery. In interprofessional focus groups, pediatric clinicians emphasized the importance of considering safety and privacy to promote rapport and *webside manner* when conducting virtual visits. The inclusion of *webside manner* instructions within training curricula is crucial as telemedicine becomes an established modality for providing health care.

(JMIR Hum Factors 2021;8(4):e29941) doi:10.2196/29941

KEYWORDS

pediatrics; telemedicine; video visits; communication; webside manner

Introduction

Background

Telehealth is a broad term that describes the provision of health care remotely using technological tools with or without a video connection [1]. Telemedicine is a subset of telehealth that refers specifically to the provision of *clinical* health care services. This can involve asynchronous transmission (ie, store and forward) of information for later review by a clinician or synchronous, live conferencing [2]. In a virtual visit, the patient and clinician are connected via a live, synchronous, interactive video system.

Within pediatrics, before the onset of the COVID-19 pandemic, telemedicine had been used in a variety of specialties, including neonatology, critical care, ophthalmology, dermatology, and urology [3-6]. Although there are additional concerns and logistical issues in implementing this type of care delivery in pediatrics, several studies have also demonstrated that the use of this technology is feasible, safe, economical, and beneficial to families because of reduced absenteeism for children and their caregivers [4,7,8]. Within this context, a telemedicine program was launched at our institution, a large tertiary care pediatric center, in late 2016.

Although telemedicine was originally used to access patients in remote locations, virtual visits have increasingly been accepted as a tool to provide real-time, convenient medical care. In part, this is because of the rapid advances in technology and the widespread affordability and accessibility of basic telemedicine tools (eg, mobile devices) [9]. Since its inception in late 2016, our institutional telemedicine program rapidly expanded to include 22 departments and 2345 virtual visits (<1% of total outpatient visits) when the study was initiated in November 2018. Patients participated in the virtual visit from their home or any other convenient location. The virtual visit is the telemedicine focus in this analysis.

Objectives

With the rapid application of this innovative technology, it is essential to preserve standards for high-quality and meaningful care. This includes effective virtual communication or good *webside manner* [10,11]. When communicating through technology, bedside manner, or the way in which clinicians relate with patients, may not be implemented in the same way as in person. Although maintaining a connection with patients and having therapeutic in-person interactions are considered *good* bedside manner, the ability to connect with patients virtually or in a webside manner is a novel concept. Modifications to the environment and clinician communication style may be necessary to build rapport and positively affect the visit experience. As virtual visits are becoming a powerful tool for clinicians to connect with children and families, it is essential that clinicians develop these skills. The aim of this study, conducted before the COVID-19 pandemic, is to understand the experiences of telemedicine providers with patient-to-provider virtual visits and communication with families at a single large-volume children's hospital. We anticipate that these qualitative data will be useful for guiding program development and training future clinicians, information that remains salient, given the established role of virtual visits in the context of the COVID-19 pandemic.

Methods

Study Design

A total of 2 focus groups of clinicians performing virtual visits were conducted before the COVID-19 pandemic, with the aim of generating discussions around shared experiences in the implementation and practice of telemedicine in their individual disciplines. The institutional review board of Boston Children's Hospital deemed this study exempt.

Sample

A purposive sample of clinicians who performed virtual visits at a single pediatric institution was recruited to participate in the study [12]. In keeping with purposive sampling, potential participants were invited based on a desire to represent groups that were already more extensively involved in virtual visits (clinical champions), becoming more involved in virtual visits (those increasing their volume), and representing both medical and surgical specialties as well as different health care disciplines (Doctor of Medicine vs nurse practitioner, registered nurse, or physician assistant). The participants did not need a minimum number of years of experience with virtual visits, and not all clinicians were contacted. Dedicated clinical champions were working with the hospital's virtual visit team to increase virtual visit volume and engagement in their department.

To reflect the diversity of experiences with virtual visits at the institution, groups were constructed based on specialty (both medical and surgical) and telemedicine experience (defined as the number of virtual visits conducted). At the time of the study, because of insurance restrictions, virtual visits at the institution were limited primarily to postoperative and established visits. Recruitment for the first focus group included dedicated clinical champions for the virtual visit team, and the second group included clinicians who actively increased their virtual visit volume. Clinicians were contacted via email and invited to participate in the focus groups. If willing, the clinicians completed a survey and participated in a focus group. A catered breakfast was provided to the participants.

Interview Guide and Procedures

On the basis of a review of the literature, interprofessional collaboration, and discussion with clinicians with telemedicine experience, a clinician survey and focus group guide were developed. The survey contained a mix of 5-point Likert scales, binary (yes or no), and multiple-choice questions that covered topics including the individual clinician experience with virtual visits and their opinions on the efficacy of the visits. The semistructured focus group guide contained open-ended questions regarding clinicians' general experience with virtual visits, impression of preparedness, strategies for establishing rapport with patients and families, and suggestions for future considerations (Multimedia Appendix 1).

The 2 focus groups were conducted to include clinicians from the following disciplines: primary care, urology, ophthalmology, gynecology, cardiac surgery, psychiatry, neurosurgery, and

orthopedics. Groups were moderated by 1 team member and supported by 3 others. One of the team members was a virtual visit clinical champion and participated in the first focus group and served as an observer in the second. Clinicians participated in person or by phone, and all sessions were digitally recorded and professionally transcribed.

Data Analysis

Field notes and focus group transcripts were reviewed to identify themes that clinicians voiced on experiences related to their virtual visit encounters. Thematic analysis is a common research strategy among qualitative researchers. It enables investigators to form generalized commentary in a subject area via the compilation of participant-level experiences and opinions [13,14]. After review and discussion, an initial set of codes was generated by the team and then applied to the transcript data. Meetings were then held to resolve any discrepancies, and a final coding framework was agreed upon. The analysis generated a clear saturation of thematic content between the 2 focus groups. After the iterative coding process was complete, the research team used NVivo 12 software (QSR International) to organize the data for further discussion. The qualitative data were then iteratively reviewed so that codes could be collated into themes and subthemes. Through this process, 3 overarching themes were identified that best described and compiled the body of data.

Results

Participant Characteristics and Survey Results

The focus groups were conducted in November 2018 and March 2019. A total of 11 clinicians participated in the study, who were split between the 2 focus groups of 6 (55%) and 5 (45%) participants. Approximately 73% (8/11) were physicians, and the groups were divided into medical and surgical specialties (Table 1). The focus group duration was an average of 69 (SD ± 8.5; range 63-75) minutes.

The survey results from the participants (8/11, 73%) are summarized in Table 2. All respondents answered each question.

Table 1. Focus group participant characteristics (N=11).

Participants	Focus group 1 (n=6)	Focus group 2 (n=5)
Number of participants by subspecialty type, n (%)		
Surgical	3 (50)	4 (80)
Medical	3 (50)	1 (20)
Number of participants by clinician type, n (%)		
MD[a]	4 (67)	4 (80)
PA[b], RN[c], NP[d]	2 (33)	1 (20)
Number of virtual visits completed, range[e]		
MD participants	1-106	4-135

[a]MD: Doctor of Medicine.

[b]PA: physician assistant.

[c]RN: registered nurse.

[d]NP: nurse practitioner.

[e]Data for physician assistant, registered nurse, and nurse practitioner participants were not available.

Table 2. Clinician survey results.

Question	Answer options	Responses, median (IQR)
Approximately how many virtual visits have you completed?	Free text	50 (7.25-60.5)
How prepared did you feel to start virtual visits after virtual visit training?	1. Not at all 2. Not really 3. Neutral/I don't know 4. A little bit 5. Completely	4.00 (3.75-4.25)
Generally, how satisfied have you been with the virtual visit experience?	1. Not satisfied 2. Slightly satisfied 3. Neutral/I don't know 4. Very satisfied 5. Extremely satisfied	4.00 (4.00-4.25)
What measures do you take to minimize background noise or change other environmental conditions that may affect the quality of the encounter?	1. Wearing headphones 2. Conducting the virtual visit in a private office or space 3. Put a sign on my door 4. Use partitioning wall 5. Other, please specify_____	Conducting the virtual visit in a private office or space: 5 clinicians; Put a sign on my door: 1 clinician; Other: made sure I had a wall or normal plane behind me: 1 clinician
I am able to communicate effectively with the patient and family.	1. Strongly disagree 2. Disagree 3. Neutral/I don't know 4. Agree 5. Strongly agree	4.5 (4.00-5.00)
I am able to obtain sufficient information even though the physical examination is not in-person.	1. Strongly disagree 2. Disagree 3. Neutral/I don't know 4. Agree 5. Strongly agree	4.00 (3.25-4.00)
How do you see providers being educated on virtual visits in the future?	1. In-person training 2. Self-paced online learning 3. Interactive simulation 4. Other, please specify _____	In-person training: 5 clinicians; Self-paced online learning: 3 clinicians; Interactive simulation: 6 clinicians
By performing virtual visits, I am able to offload in-person visits.	1. Strongly disagree 2. Disagree 3. Neutral/I don't know 4. Agree 5. Strongly agree	4.5 (4.00-5.00)

Thematic Analysis

A model emerged from the analysis that contained 3 overarching themes: (1) development of best practices, (2) barriers to scalability, and (3) establishing patient rapport. The generation of these themes suggested their applicability across participants from different disciplines, although individual clinician experiences were influenced by their subspecialty and longevity of virtual practice.

Theme 1: Development of Best Practices

Overview

Overarching the discussions was a need to develop best practices in pediatric telemedicine, including but not limited to the need to determine the ideal virtual patient, address privacy concerns, and ensure adequate physical examinations. Clinicians agreed that different disciplines could learn from one another and that, although some issues cut across disciplines, others were unique to individual subspecialties.

Ideal Patient

The ideal telemedicine patient was described by focus group participants as one who would experience the potential benefits of telemedicine (ie, living far from the hospital and whose parents' capacity to take time off work was limited), is already comfortable with the clinician, and whose physical evaluation requires minimal hands-on examination. Multiple clinicians shared that it was important to thoughtfully select the patients who would gain the most value from the telemedicine experience. Clinicians who represented mental health fields additionally expressed that some patients did measurably *better* with virtual visits than with in-person visits:

We're doing these appointments not because we expect a postoperative complication because we've kind of screened these patients out, but we're doing them as a touch point to the patient so that they feel cared for and so that their perception of care is better just because we're looking at them and we're talking to them.

We also have a lot of patients that are coming from like South Shore or just a long ways away from the hospital, I think like all of us and it's so nice to be able to...especially if it's a visit where we're kind of just checking in on like their experience on a medication or side-effects where it may only need to be like a 20-minute conversation that can happen without a two-hour drive. That feels good for everybody so I don't think there's been any downside to it that we've seen yet.

There's a number of kids, if they are on the autism spectrum or have connection difficulties socially, it seems that like I've had communication is easier for them, like it's more approachable for them. And so sometimes we're just able to get more out of them than we would if we were in person where there's something that is physically just...in the room it's hard for them about connecting in person.

Privacy Concerns

Across disciplines, clinicians shared concerns regarding patient privacy and exchanged best practices for dealing with sensitive physical examinations:

I've started telling patients, just because this was on my mind about the privacy issues, etc., their comfort level, I started telling patients that I'm in my office, I'm in my private office, and nobody is going to open the door.

I do say that in the office for the older kids...we're going to examine down here. You only do this if there's a doctor and your parents in the room. I guess I really hadn't said that when I'm on the telephone...that's probably a good idea.

We also built in the support piece so probably we wanted to make sure that the patient felt comfortable and safe and so we have a social worker call them right after the visit, I contact them a week later just to make sure that they're feeling okay with it.

Physical Examination

Regarding the virtual physical examination, most clinicians felt they could do a *good enough* examination for the purpose of the visit. Some clinicians expressed concerns regarding the patient's or parent's impression of the examination, often because of the technical issues with video equipment or lack of user experience. Clinicians also shared how they adapted their use of technology to meet the needs of their specific clinical practice:

It only took probably a handful of cases to make it obvious that we do see what we need to see very clearly. I think it's exceeded all of our expectations for sure.

Sometimes it does lead to maybe a suboptimal exam where you're like okay, well, I'm sure it's fine, I can see it well enough but I'd love to see it better and it's certainly not the same as seeing them in person; it's just probably "good enough." But I love leaving those with thinking this virtual visit was equivalent to my physical, to my in-person physical examination and often they do feel that way. But when the camera is jiggly or the connection's not great, I don't feel that way. I feel that it's good enough, but that's a little bit of a slippery slope if you think about it.

So yesterday I had the big sister actually get on the other side of the iPad and have the baby look at the big sister and then I said okay, I see the eyes go to the left, so I had like sister, big sister run to the left and the baby goes ooh, follows the kid, and I'm like mom, hold the head so I can see the eyes moving and it was an awesome way to do the exam. And so, then the big sister was running back and forth and the kid's eyes are going back and forth, I'm like this is great.

Theme 2: Barriers to Scalability

Overview

Many clinicians mentioned the challenges they faced, which made them concerned about the quality and effectiveness of virtual visits. Some issues involved overcoming technical difficulties for the patients, families and clinicians, which occupied time during the visit. There were also concerns regarding privacy and how that might limit the environment in which a visit could take place.

Family Preparation and Education

Clinicians noted that although previsit educational materials were provided to families, these materials were not adequate, as clinicians spent a significant amount of time assisting patients and families with technical issues:

I estimate like probably 20% of my time, of my patient load is spent doing a lot of explaining of things.

I find that they're not reading these things and so we're trying to educate them but our education has not been effective thus far with the handouts, with the carousel screen.

That's my problem. More than 50% of my visits are spent with 50% of my time teaching best practices and the more that I've realized that the best practices actually allow me to see that postoperative surgical wound better, to give that equivalent virtual experience as the in-person experience, the more frustrated I've become with the idea that the patients I don't think are reading these best practices.

Privacy and Safety

Clinicians expressed concern that patients and families did not receive adequate information regarding the privacy and confidentiality of encounters. Some suggested changes that

could be made individually and at a program level to improve patient and family knowledge of privacy restrictions around virtual visits. Clinicians expressed apprehension regarding performing sensitive examinations virtually. This was especially concerning for specialties such as urology, in which the physical examination primarily involves sensitive areas of the body:

> *And so to this point of preparation, we have no sort of documentation; we have no pre-visit preparation telling them these things, that your provider will be in a private spot, there will be nobody else present, etc., your privacy is guarded, we have no...guidance on this with the families at all. I think we have to be much more careful with all these things to make sure that they come away feeling really confident and safe in this experience.*
>
> *So at this point the patients and the families don't receive anything that says that you can feel safe and that any pictures obtained during the...there's nothing like that?*
>
> *Well couldn't it be part of the carousel, even just a reminder that this is still a private appointment and that any information obtained is really part of your health record, that would be just to remind them because they have to read those things because they go around.*
>
> *So I don't know, it's something that I think for those that do that sort of sensitive exam to really consider how we should best prepare families to do that*
>
> *And these are interactions with these patients virtually with sensitive exams—we have not set any expectations, we have not set any boundaries.*

Logistical Issues

Clinicians acknowledged issues with patients and families using technology that limited the effectiveness and impact of virtual interactions. Clinicians expressed frustration with the technology not functioning as well as they thought it should, resulting in a poor connection or an image that would make the visit difficult:

> *…it's a struggle, to get—to teach them how to turn, you know, how to reverse the camera and they're like what, and then they hang up*
>
> *...sometimes physical exam is—It's impossible either A, you can't figure out how to focus the camera, kids moving. Their whatever, bandwidth is horrible, so you're—it's like this blurred image anyway. A lot of it they just can't get it and eventually you're just like okay, good enough. That's it.*
>
> *...the issue is like the camera is in the corner but the image is in the center of the screen. So, like if this is the baby, getting them to line up the corner with what you want to see, versus the middle*
>
> *...but it's just the optics that whatever the bandwidth is not always good enough that we can actually see enough...Nothing you can do about it. Either I have to bring them back in...Decide on how important it is.*
>
> *If the bandwidth is bad there's just nothing—you're going to have to bring them in eventually. That's probably the biggest limitation at the moment for me.*
>
> *Well, I understand that there's only so many limitations, I can't call the help desk and say, hey guess what, their WI-FI is horrible*

Theme 3: Establishing Patient Rapport

Overview

Clinicians acknowledged a learning curve in their ability to use telemedicine technology to provide optimal patient experience. Part of this is based on their own comfort level and confidence but is complicated by learning how to establish new ways to interact with and build rapport with families via this modality. As such, not all clinicians come away with a positive impression of the experience.

Clinician Confidence and Flexibility

Clinicians discussed how their own comfort, confidence, and flexibility were critical to the effective use of virtual visits:

> *I think it's getting more comfortable behind the camera; it's just I think being less stiff and sort of bringing what I bring to the bedside to the camera and in trying to remember that and not being uncomfortable with the media part of this.*
>
> *They have to understand that I believe in it and that it's working. Like in the beginning, because I wasn't sure myself and I had to figure it out, and how am I going to talk to them about it and all this, and I no longer say this is something we're trialing out and all that. You know, we haven't done it before. I just say this is what we do. You'll find it very helpful.*

Patient Interactions

Clinicians acknowledged the importance of establishing rapport with families via virtual visits. They noted that interactions can be less natural and expressed particular anxieties about certain circumstances that virtual visits may make more challenging, such as communicating about privacy issues:

> *And sort of getting that whole thing in there and I think connecting in a personal way, for me, is acknowledging some of the difficulties, especially for our population and this is just such a huge undertaking for parents and families and just saying kind of hang in there, that kind of thing. So I think those personal statements from me are important*
>
> *...it's been a little weird, I've got to be honest, with little boys that are old enough, the mom calls them over and what happens in the office is the parents will routinely say "remember, Jimmy, only mommy, daddy and the doctor," right, and then I say "exactly," talk them through it, make sure they're comfortable. On the video it's mom saying "okay, Jimmy, remember it's only video when it's the doctor"*

Clinicians' Impression of the Virtual Interaction

Clinicians had both positive and negative impressions of the virtual interaction and their ability to establish rapport with

families via virtual visits. Clinicians seemed surprised by their positive experiences. Negative experiences focused on control of the encounter:

> *I just feel like...that I have an ability to still connect with patients the way that I like to. I still take in information the way I do as a clinician. I do look around the room, I do look at siblings, I do look at the interaction of parents, those kinds of things, so it's not an isolated FaceTime experience I think that you would...and I expected it to be kind of a little bit sterile or super-removed I think.*

> *I have found that some of my patients have used this as a liberty to change the way this relationship is going to work, that all of a sudden now all the kids can run around, all of a sudden like other things can go on like the plumber coming in to fix the house at the same time as our visit. Now I know some of these things are out of control, like I get a page in the middle of a visit and I have to step out of an appointment, so it happens, but my impression is that the percentage of times that this happens is higher when they're in the home environment versus when they are in our office environment.*

Discussion

Principal Findings

This qualitative study of clinicians at a large, academic pediatric medical center who were initial users of telemedicine (before the COVID-19 pandemic) identified 3 key themes that are valuable to the understanding of how patient-to-provider virtual visit programs may be sustainable and generalizable for pediatric patient care in the long term. These included a need to develop best practices in pediatric telemedicine, particularly regarding patient selection, privacy and physical examination; barriers to scalability, including technical and logistical issues as well as privacy concerns; and the ability of clinicians to establish rapport with patients through virtual visits. Issues and concerns related to privacy were salient across all themes, and clinicians noted opportunities for shared learning across subspecialties.

The last 2 decades have seen a growth in the use of telemedicine, particularly for medically underserved communities [9,15-18]. Advocates for its use in pediatrics have pushed to reduce barriers as a means of expanding access to pediatric care [9]. However, addressing the need for stable funding and adequate training in technology use has been highlighted as an important priority for the forward expansion of telemedicine [9], and this study makes it apparent how essential this is for the successful implementation of telemedicine in the pediatric setting. As health care clinicians learn their clinical skills, part of their training is the development of tools and habits they will use to establish positive bonds with their patients. The clinician's ability to interact with patients and deliver high-quality care can be described as bedside manner. The concept of *webside manner* was introduced to highlight that telemedicine interactions may require new training and learned skills in a variety of domains, including technology, to ensure the same level of clinician interaction with patients [10,19]. This qualitative study of telemedicine providers brings to light the importance of formal training in *webside manner* to optimize the virtual visit experience.

Comparison With Prior Work

This study reflected the views of initial telemedicine users before the COVID-19 pandemic. The findings were comparable with a prior framework of early adopters, which highlights the need for clinicians to be flexible and attentive to the nonmedical aspects of patient interaction [20]. Despite the prevalence of technology in health care, clinicians in this study reported technical and logistical issues that affected the virtual visits. Families were given instructions on how to use telemedicine; however, many still had trouble or did not fully understand the instructions. In addition, some clinicians raised concerns that the video quality may not be adequate for all situations.

Similar concerns about technology and patient selection have been raised elsewhere in the literature [15,21-23]. A qualitative study among rural health clinicians in the United States also specifically identified concerns about how technology may affect personal relationships, in this case among generalists and subspecialists [16]. A qualitative study in Australia of rural and urban health care clinicians with variable levels of exposure to telemedicine identified that those with greater telemedicine experience recognized the need to be pragmatic about the risks and challenges of telemedicine as well as for ongoing technology support [24]. In this study, clinicians broadly reflected on their rapid experiential learning and had varied attitudes regarding their comfort with patient interaction. Other studies evaluating clinician attitudes around telemedicine integration into pediatric care similarly suggest that contextual factors, such as perceived usefulness of telemedicine and ease of use, may affect uptake and concern regarding the impact of technology on the patient relationship [16,25].

The 2 other prominent issues in focus group discussions were safety and privacy. Traditional health care visits take place in controlled environments that are designed to ensure safety and privacy, allowing the clinician to focus on the patient and their family. In a virtual visit, the clinician has limited control over where the patient and family is during the visit, and the patient cannot see the clinician's surroundings. Consequently, privacy is a concern raised by both clinicians and patients, and adequate education and tools are of paramount importance [15,23,26,27]. Clinicians in this study agreed on the importance of ensuring safety and privacy, although the methods and tools used to address them varied. During the focus groups, clinicians had a real-time exchange of ideas regarding their telemedicine improvement strategies, and their engagement in learning from one another around this issue supports the need for further development of best practices in this area. Some actionable recommendations that follow from these discussions are outlined in Table 3.

Table 3. Actionable recommendations.

Category	Actionable recommendation	Responsible for implementation
Technology	• Incorporate tools to support previsit technical testing (eg, video, audio, and connection) for both patients and clinicians • Develop and use HIPAA[a]-compliant methods for patient-to-provider sharing of content (eg, photos and laboratory data) before or during the virtual visit	Institution
Environment	• Ensure that the physical environment supports a private and professional virtual encounter	Clinician
Training	• Create standardized patient and clinician user guides for the virtual visit platform; include major technical issues, best practices, and explanation of privacy issues • Include specific physical exam guidance, depending on subspecialty, including language around privacy • Share written content and links to published resources for virtual visits • Include simulation for onboarding and virtual visit training with a mock, recorded virtual visit	Institution
Webside manner	• Pay attention to the nonmedical aspects of the interaction (eg, eye contact) to ensure the most favorable patient experience	Clinician

[a]HIPAA: Health Insurance Portability and Accountability Act.

Although providers in this study expressed concern about the scalability of telemedicine before the COVID-19 pandemic, the pandemic reframed thinking about the potential barriers to telemedicine, such as financial concerns about reimbursement, credentialing and licensing, and medical liability [15]. These issues have been addressed by government-mandated policies in the short term and may have some residual impact on framing access to telemedicine services going forward [15,28-32]. Some pediatric settings have demonstrated the ability to rapidly scale telemedicine during COVID-19, developing novel mechanisms to connect with families that ensure privacy [33]. These success stories offer opportunities for *proof of concept* demonstrations, whereby eliminating some of the barriers has paved the way for infrastructural scaling. This may suggest that some of the logistical concerns regarding scaling expressed by clinicians in the study may be addressed by continued advances in telemedicine technology, adequate information technology support, and ongoing relaxation in these other areas.

However, it does not fully address the concerns regarding logistical issues for patients and families and aspects of privacy that are based not solely on technology but also on how clinicians are educated to interact with families via telemedicine platforms. Consistent with other studies [34], we demonstrate that ensuring long-term success with telemedicine will require an appropriate selection of patients and education for clinicians as well as patients and families. Furthermore, despite the recent need and swift implementation of telemedicine, this study's findings, reflecting a range of specialties and professional backgrounds, suggest that enduring concerns regarding selecting the medical conditions or circumstances for which telemedicine is appropriate, privacy concerns, and the impact on the patient–clinician interaction will warrant ongoing attention to ensure that access is adequately balanced with quality [15,19,34-36]. Given the rapid expansion of telemedicine because of the COVID-19 pandemic, recognizing how to address some of these concerns is critical to ensuring that a broader range of clinician and patient experiences are optimized.

Strengths and Limitations

This study has a number of strengths. All participants were pediatric clinicians, which provided internal consistency. Specific issues of relevance to the pediatric visit included the participation of multiple family members, control of the technology by someone other than the patient, and the challenge of managing pediatric comprehension of technology. In addition, the study was completed before the COVID-19 pandemic, which has increased telemedicine use exponentially. As such, the clinicians in our study had more control over their early telemedicine practice and the ability to choose which patients were the most appropriate for virtual visits. Therefore, the perspectives of these clinicians provide insight into the implementation lessons of telemedicine before the COVID-19 pandemic, when scaling and coverage became the priority.

Several limitations of this study should be noted. The study took place at a single pediatric institution in Massachusetts, which at the time had some of the more restrictive laws regarding telemedicine. The reimbursement structure limited the types of patients and visits that could be performed. Accordingly, most clinicians were from surgical specialties, and many of the visits being conducted were postoperative visits, as these fall under the global charge capture. However, the intent was not that study findings would reflect the full range of practice available by virtual visits but rather inform the perspectives of participants and future directions. It should be noted that one of the authors participated in the first focus group. As this author was not involved in data analysis and there were similar themes in both focus groups, we are confident that this did not introduce bias into the discourse. Finally, the focus of this study was on clinician perspective. Although not specifically a limitation of the methods, exploring the patient and family perspective of telemedicine should be an important component of future studies and can be informed by the themes identified here. In addition, further investigation with other provider groups with different types of telemedicine experience would strengthen the potential for generalizability of the study findings.

Conclusions

It is likely that a substantial portion of clinical practice will continue to be performed virtually even after the COVID-19 pandemic. The findings of this study suggest that telemedicine curricula, including instruction on *webside manner* skills, should be incorporated into medical training. Integration of this training for all programs will be crucial to the efficacy and sustainability of this essential mode of health care delivery. Some medical school programs have already taken on this challenge [37]. As one focus group participant noted, "I think that the broader concept of webside manner should apply to all of us, kind of like what we do in medical school, PA school, nursing school, so that would be my hope in terms of future direction." Future work may address how such training affects the patient–clinician experience and the ability to further scale telemedicine.

Acknowledgments

EST was supported by grant numbers T32HS000063 from the Agency for Healthcare Research and Quality and K12DK094721 from the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality or the National Institutes of Health. MVC was supported by the National Library of Medicine Biomedical Informatics and Data Science Research Training grant T15LM007092.

Authors' Contributions

JBF, MVC, ASS, AFC, CS, and CB conceptualized the study, designed the focus group guide, and conducted the focus groups. JBF, EST, MVC, AFC, and ETR participated in data analysis and interpretation and drafted the initial manuscript. All authors reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Conflicts of Interest

ETR was the site principal investigator for a clinical trial sponsored by Astra Zeneca. The other authors have no financial relationships or potential conflicts of interest relevant to this paper to disclose.

Multimedia Appendix 1
Focus group guide discussion questions.
[DOCX File , 15 KB - humanfactors_v8i4e29941_app1.docx]

References

1. Dorsey ER, Topol EJ. State of telehealth. N Engl J Med 2016 Jul 14;375(2):154-161. [doi: 10.1056/NEJMra1601705] [Medline: 27410924]
2. The American Telemedicine Association. URL: https://www.americantelemed.org/ [accessed 2020-05-25]
3. Burke BL, Hall RW, SECTION ON TELEHEALTH CARE. Telemedicine: pediatric applications. Pediatrics 2015 Jul;136(1):e293-e308 [FREE Full text] [doi: 10.1542/peds.2015-1517] [Medline: 26122813]
4. Finkelstein JB, Cahill D, Young K, Humphrey K, Campbell J, Schumann C, et al. Telemedicine for pediatric urological postoperative care is safe, convenient and economical. J Urol 2020 Jul;204(1):144-148. [doi: 10.1097/JU.0000000000000750] [Medline: 31951495]
5. Saleem SM, Pasquale LR, Sidoti PA, Tsai JC. Virtual ophthalmology: telemedicine in a COVID-19 era. Am J Ophthalmol 2020 Aug;216:237-242 [FREE Full text] [doi: 10.1016/j.ajo.2020.04.029] [Medline: 32360862]
6. Bowe T, Hunter DG, Mantagos IS, Kazlas M, Jastrzembski BG, Gaier ED, et al. Virtual visits in ophthalmology: timely advice for implementation during the COVID-19 public health crisis. Telemed J E Health 2020 Sep;26(9):1113-1117. [doi: 10.1089/tmj.2020.0121] [Medline: 32408801]
7. McConnochie KM, Wood NE, Kitzman HJ, Herendeen NE, Roy J, Roghmann KJ. Telemedicine reduces absence resulting from illness in urban child care: evaluation of an innovation. Pediatrics 2005 May;115(5):1273-1282. [doi: 10.1542/peds.2004-0335] [Medline: 15867035]
8. Lesher AP, Shah SR. Telemedicine in the perioperative experience. Semin Pediatr Surg 2018 Apr;27(2):102-106. [doi: 10.1053/j.sempedsurg.2018.02.007] [Medline: 29548350]
9. COMMITTEE ON PEDIATRIC WORKFORCE, Marcin JP, Rimsza ME, Moskowitz WB. The use of telemedicine to address access and physician workforce shortages. Pediatrics 2015 Jul;136(1):202-209 [FREE Full text] [doi: 10.1542/peds.2015-1253] [Medline: 26122802]
10. Gonzalez R. Telemedicine is forcing doctors to learn webside manner. WIRED. 2017. URL: https://www.wired.com/story/telemedicine-is-forcing-doctors-to-learn-webside-manner/ [accessed 2020-05-15]
11. Teichert E. At the 'webside'. Mod Healthc 2016 Oct;46(35):16-18. [Medline: 30452797]
12. Rice PL, Ezzy D. Qualitative Research Methods: A Health Focus. Oxford: Oxford University Press; 1999:1-295.

13. Boyatzis RE. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, CA: Sage Publications Inc; 1998:1-200.
14. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
15. Utidjian L, Abramson E. Pediatric telehealth: opportunities and challenges. Pediatr Clin North Am 2016 Apr;63(2):367-378. [doi: 10.1016/j.pcl.2015.11.006] [Medline: 27017042]
16. Ray KN, Demirci JR, Bogen DL, Mehrotra A, Miller E. Optimizing telehealth strategies for subspecialty care: recommendations from rural pediatricians. Telemed J E Health 2015 Aug;21(8):622-629 [FREE Full text] [doi: 10.1089/tmj.2014.0186] [Medline: 25919585]
17. Davis MM, Currey JM, Howk S, DeSordi MR, Boise L, Fagnan LJ, et al. A qualitative study of rural primary care clinician views on remote monitoring technologies. J Rural Health 2014;30(1):69-78 [FREE Full text] [doi: 10.1111/jrh.12027] [Medline: 24383486]
18. Marcin JP, Ellis J, Mawis R, Nagrampa E, Nesbitt TS, Dimand RJ. Using telemedicine to provide pediatric subspecialty care to children with special health care needs in an underserved rural community. Pediatrics 2004 Jan;113(1 Pt 1):1-6. [doi: 10.1542/peds.113.1.1] [Medline: 14702439]
19. McConnochie KM. Webside manner: a key to high-quality primary care telemedicine for all. Telemed J E Health 2019 Nov;25(11):1007-1011. [doi: 10.1089/tmj.2018.0274] [Medline: 30648924]
20. Waller M, Taylor L, Portnoy J. The medical virtualist: is pediatric patient care using telemedicine, a new specialty? Pediatr Ann 2019 Jun 01;48(6):e243-e248. [doi: 10.3928/19382359-20190520-01] [Medline: 31185116]
21. Moffatt JJ, Eley DS. Barriers to the up-take of telemedicine in Australia--a view from providers. Rur Remot Health 2011;11(2):1581 [FREE Full text] [Medline: 21385004]
22. McConnochie K, Wood N, Herendeen N, ten Hoopen C, Denk L, Neuderfer J. Integrating telemedicine in urban pediatric primary care: provider perspectives and performance. Telemed J E Health 2010 Apr;16(3):280-288 [FREE Full text] [doi: 10.1089/tmj.2009.0112] [Medline: 20406114]
23. McSwain SD, Bernard J, Burke BL, Cole SL, Dharmar M, Hall-Barrow J, et al. American Telemedicine Association operating procedures for pediatric telehealth. Telemed J E Health 2017 Sep;23(9):699-706. [doi: 10.1089/tmj.2017.0176] [Medline: 28829680]
24. Shulver W, Killington M, Crotty M. 'Massive potential' or 'safety risk'? Health worker views on telehealth in the care of older people and implications for successful normalization. BMC Med Inform Decis Mak 2016 Oct 12;16(1):131 [FREE Full text] [doi: 10.1186/s12911-016-0373-5] [Medline: 27733195]
25. Ray KN, Felmet KA, Hamilton MF, Kuza CC, Saladino RA, Schultz BR, et al. Clinician attitudes toward adoption of pediatric emergency telemedicine in rural hospitals. Pediatr Emerg Care 2017 Apr;33(4):250-257. [doi: 10.1097/PEC.0000000000000583] [Medline: 26785087]
26. Kuziemsky CE, Gogia SB, Househ M, Petersen C, Basu A. Balancing health information exchange and privacy governance from a patient-centred connected health and telehealth perspective. Yearb Med Inform 2018 Aug;27(1):48-54 [FREE Full text] [doi: 10.1055/s-0038-1641195] [Medline: 29681043]
27. Ray KN, Ashcraft LE, Mehrotra A, Miller E, Kahn JM. Family perspectives on telemedicine for pediatric subspecialty care. Telemed J E Health 2017 Oct;23(10):852-862 [FREE Full text] [doi: 10.1089/tmj.2016.0236] [Medline: 28430021]
28. March 15 2020 Telehealth Order. Commonwealth of Massachusetts. URL: https://www.mass.gov/doc/march-15-2020-telehealth-order [accessed 2020-05-08]
29. Medicare telemedicine health care provider fact sheet. Centers for Medicare & Medicaid Services. 2020 Mar 17. URL: https://www.cms.gov/newsroom/fact-sheets/medicare-telemedicine-health-care-provider-fact-sheet [accessed 2020-05-08]
30. COVID-19 Telehealth Coverage Policies. Center for Connected Health Policy. URL: https://www.cchpca.org/resources/covid-19-telehealth-coverage-policies [accessed 2020-05-08]
31. Gutierrez J, Kuperman E, Kaboli PJ. Using telehealth as a tool for rural hospitals in the COVID-19 pandemic response. J Rural Health 2021 Jan;37(1):161-164 [FREE Full text] [doi: 10.1111/jrh.12443] [Medline: 32277777]
32. Nagata JM. Rapid scale-up of telehealth during the COVID-19 pandemic and implications for subspecialty care in rural areas. J Rural Health 2021 Jan;37(1):145. [doi: 10.1111/jrh.12433] [Medline: 32246490]
33. Patel PD, Cobb J, Wright D, Turer R, Jordan T, Humphrey A, et al. Rapid development of telehealth capabilities within pediatric patient portal infrastructure for COVID-19 care: barriers, solutions, results. J Am Med Inform Assoc 2020 Jul 1;27(7):1116-1120 [FREE Full text] [doi: 10.1093/jamia/ocaa065] [Medline: 32302395]
34. Loeb AE, Rao SS, Ficke JR, Morris CD, Riley LH, Levin AS. Departmental experience and lessons learned with accelerated introduction of telemedicine during the COVID-19 crisis. J Am Acad Orthop Surg 2020 Jun 01;28(11):e469-e476 [FREE Full text] [doi: 10.5435/JAAOS-D-20-00380] [Medline: 32301818]
35. Finkelstein JB, Nelson CP, Estrada CR. "Ramping up telemedicine in pediatric urology- Tips for using a new modality". J Pediatr Urol 2020 Jun;16(3):288-289. [doi: 10.1016/j.jpurol.2020.04.010] [Medline: 32327290]
36. Uscher-Pines L, Kahn JM. Barriers and facilitators to pediatric emergency telemedicine in the United States. Telemed J E Health 2014 Nov;20(11):990-996 [FREE Full text] [doi: 10.1089/tmj.2014.0015] [Medline: 25238565]

37. Buckley MR. The Mother of Invention: COVID-19 pandemic spurs creation of new, remote teaching methods at HMS. Harvard Medical School. 2020 May 08. URL: https://hms.harvard.edu/news/mother-invention?utm_source=Silverpop&utm_medium=email&utm_term=field_news_item_3&utm_content=HMNews05182020 [accessed 2020-05-25]

Edited by A Kushniruk; submitted 26.04.21; peer-reviewed by D Louder, M Gagnon, S Ng; comments to author 10.07.21; revised version received 05.08.21; accepted 31.08.21; published 02.12.21.

Please cite as:
Finkelstein JB, Tremblay ES, Van Cain M, Farber-Chen A, Schumann C, Brown C, Shah AS, Rhodes ET
Pediatric Clinicians' Use of Telemedicine: Qualitative Interview Study
JMIR Hum Factors 2021;8(4):e29941
URL: https://humanfactors.jmir.org/2021/4/e29941
doi:10.2196/29941
PMID:34860669

Original Paper

A Shared Decision-making Tool for Drug Interactions Between Warfarin and Nonsteroidal Anti-inflammatory Drugs: Design and Usability Study

Thomas J Reese[1], PharmD, PhD; Guilherme Del Fiol[2], MD, PhD; Keaton Morgan[2], MD; Jason T Hurwitz[3], PhD; Kensaku Kawamoto[2], MHS, MD, PhD; Ainhoa Gomez-Lumbreras[2], MD, PhD; Mary L Brown[3], PhD; Henrik Thiess[4], BSc; Sara R Vazquez[2], PharmD; Scott D Nelson[1], MS, PharmD; Richard Boyce[5], PhD; Daniel Malone[2], RPH, PhD

[1]Vanderbilt University, Nashville, TN, United States

[2]University of Utah, Salt Lake City, UT, United States

[3]University of Arizona, Tuscon, AZ, United States

[4]University of Heidelberg, Heidelberg, Germany

[5]Department of Biomedical Informatics, University of Pittsburgh, Pittsburg, PA, United States

Corresponding Author:
Thomas J Reese, PharmD, PhD
Vanderbilt University
2525 West End Avenue, Suite 1475
Nashville, TN, 37203
United States
Phone: 1 (615) 936 6867
Fax: 1 (615) 936 0102
Email: Thomas.Reese@vumc.org

Abstract

Background: Exposure to life-threatening drug-drug interactions (DDIs) occurs despite the widespread use of clinical decision support. The DDI between warfarin and nonsteroidal anti-inflammatory drugs is common and potentially life-threatening. Patients can play a substantial role in preventing harm from DDIs; however, the current model for DDI decision-making is clinician centric.

Objective: This study aims to design and study the usability of DDInteract, a tool to support shared decision-making (SDM) between a patient and provider for the DDI between warfarin and nonsteroidal anti-inflammatory drugs.

Methods: We used an SDM framework and user-centered design methods to guide the design and usability of DDInteract—an SDM electronic health record app to prevent harm from clinically significant DDIs. The design involved iterative prototypes, qualitative feedback from stakeholders, and a heuristic evaluation. The usability evaluation included patients and clinicians. Patients participated in a simulated SDM discussion using clinical vignettes. Clinicians were asked to complete eight tasks using DDInteract and to assess the tool using a survey adapted from the System Usability Scale.

Results: The designed DDInteract prototype includes the following features: a patient-specific risk profile, dynamic risk icon array, patient education section, and treatment decision tree. A total of 4 patients and 11 clinicians participated in the usability study. After an SDM session where patients and clinicians review the tool concurrently, patients generally favored pain treatments with less risk of gastrointestinal bleeding. Clinicians successfully completed the tasks with a mean of 144 (SD 74) seconds and rated the usability of DDInteract as 4.32 (SD 0.52) of 5.

Conclusions: This study expands the use of SDM to DDIs. The next steps are to determine if DDInteract can improve shared decision-making quality and to implement it across health systems using interoperable technology.

(JMIR Hum Factors 2021;8(4):e28618) doi:10.2196/28618

KEYWORDS

shared decision-making; user-centered design; drug interaction; clinical decision support

Introduction

Background

Drug-drug interactions (DDIs) are preventable adverse events that are responsible for 5% to 14% of adverse drug reactions in patients that are hospitalized [1,2], are a major risk factor for hospitalization [3], and occur in up to 13% of older adult ambulatory patients [4-6]. Exposure to life-threatening DDIs occurs despite the widespread use of clinical decision support. Alarmingly, up to 24% of patients on warfarin receive a prescription for a nonsteroidal anti-inflammatory drug (NSAID), which increases the risk of gastrointestinal bleeding up to twofold [7,8].

Most electronic health records (EHRs) implement DDI clinical decision support functionality with underlying logic provided by drug knowledge base vendors, but DDI alerts continue to be overridden at rates as high as 90% [9-11]. The current model for DDI decision-making is highly clinician-centric in spite of the fact that patients can play a substantial role in preventing potential harm due to DDIs. Studies that have explored different clinical decision support for DDIs [12,13] indicate that interactive decision dashboards have the potential to foster informed decision-making by patients [13]. These decision aids allow patients and clinicians to deliberate together about the advantages and disadvantages of different therapies and arrive at decisions that are concordant with the best available evidence, clinician knowledge, and patient preferences [14,15].

Accordingly, the overarching goal of this study is to incorporate patient-centered shared decision-making (SDM) for addressing DDIs, an advance from clinician-centric decision-making models. SDM is a conversation where patients share their values and preferences to choose a treatment that aligns with their goals [16,17]. Electronic decision aids can support this conversation; however, SDM is uncharted in the DDI domain [18].

Objectives

The purpose of this study was to design and evaluate the usability of DDInteract, an SDM tool for the warfarin and NSAID DDI.

Methods

Overview

The design and usability assessment of DDInteract was guided by user-centered design principles and an SDM framework. The user-centered process included iterative and overlapping steps of prototyping (ie, low fidelity, stable, and high fidelity), stakeholder feedback, and usability heuristics and testing [19-24]. The SDM framework consists of five steps: (1) seek your patient's participation, (2) help your patient explore and compare treatment options, (3) assess your patient's values and preferences, (4) reach a decision with your patient, and (5) evaluate your patient's decision [25]. Figure 1 depicts a summary of the design and usability process. This study was approved by the University of Utah Institutional Review Board.

Figure 1. Summary of the design and usability process for DDInteract. IPDAS: International Patient Decision Aid Standards Collaboration.

Design

The design team consisted of multidisciplinary experts in DDIs, clinical decision support, patient and provider communication, SDM, and pharmacotherapy outcomes. The process began with an artifact appraisal of SDM tools, DDI alerts, and clinical practice materials relevant to anticoagulants. Information from this appraisal was used to sketch low-fidelity feature prototypes, which were reviewed and discussed with the design team in weekly meetings. Features and functionality deemed important by the design team were retained for future iterations. Features were linked to the SDM steps and checklist items from the International Patient Decision Aid Standards Collaboration [26]. Once the design team coalesced on preliminary feature designs, these were combined into an initial protype of the complete user interface using Adobe XD (Adobe Inc).

Target users (ie, 2 physicians and 1 pharmacist) were individually shown the initial complete user interface prototype and asked to provide feedback on the usefulness, aesthetics, proposed functionality, and content. Several iterations were made in collaboration with these target users until no substantial feedback was provided. At this point the prototype was considered stable enough for a heuristic evaluation. The heuristic evaluation was based on knowledge of Nielsen's 10 Usability Heuristics for User Interface Design [27,28] and was performed by two experts with training and experience in human-centered design, psychology, and medical informatics. The goal of the heuristic evaluation was to identify design flaws that could be addressed prior to conducting resource-intensive testing. Specific feedback regarding design that might impede users' goals were noted and shared in a team meeting along with a discussion of potential solutions for each flaw. Once the stable prototype was modified to address findings from the heuristic evaluation, it was considered high fidelity and ready for usability testing.

Usability

Usability assessments consisted of two parts: (1) patient interviews with simulated clinic visits and (2) clinician task performance assessments and usability surveys. DDInteract was designed to be used by clinicians at the point of care. Since patients would not use DDInteract without a clinician present, we did not test task completion success and efficiency with patients.

Patients

Patient participants were recruited from the anticoagulation service at the University of Utah. Participants were required to be on warfarin for a chronic condition such as atrial fibrillation. The perceived usability and usefulness of DDInteract was assessed with participants individually through two simulated clinical scenarios and a semistructured interview. Participants were given two short clinical vignettes to read before the session (Multimedia Appendix 1, Table S1). The decision associated with each vignette was whether to start an NSAID for pain. The vignettes were designed to test the range of responses based on a patient's risk (ie, high risk and low risk) of gastrointestinal bleeding. In the high-risk vignette, the patient had multiple risk factors for gastrointestinal bleeding including age older than 65 years, use of an antidepressant, and history of a gastrointestinal bleeding. In the low-risk vignette, warfarin was the only risk factor. Participants simulated SDM based on the clinical vignettes with a provider (author KM). Following the clinical scenarios, patients were asked questions pertaining to aspects of DDInteract, the use of DDInteract for SDM, and the utility of SDM for DDIs. The interviews were conducted online with audio and screen recording. The audio was transcribed and coded into general topics.

Clinicians

Physicians and pharmacists with anticoagulation therapy experience were recruited by snowball sampling. The overarching goal of the clinician usability assessment was to obtain objective and subjective data on the use of DDInteract. Participants were asked to complete a task performance assessment and a perceived usefulness survey. Participant characteristics were collected as part of the survey. Links to the instructional video, task performance assessment, and survey were emailed to participants. The instructional video was a brief introduction to DDInteract. The task performance assessment was web-based and recorded the participant's screen. The survey was based on the System Usability Scale and included a free-text section for feedback [29,30]. Tasks consisted of eight key navigation and functionality tasks (Table 1). Performance was measured by task completion rates and the time to complete each task. After a task was completed, the app reset to the home screen. The time was measured from when the home screen was displayed to when the task was completed.

Table 1. Clinician task prompts and actions performed that result in successful completion.

Tasks	Success
1. Your patient has questions about what a gastrointestinal bleed is. Please navigate to patient education about a gastrointestinal bleeding.	Navigating to and clicking on the drop-down arrow for "What is a gastrointestinal (Stomach) bleed?"
2. With previous patients, you have found it confusing for them to understand the drug class NSAIDs[a]. Please find the picture of multiple NSAIDs to illustrate how not only ibuprofen is an NSAID.	Navigating through the "What is a drug-drug interaction" drop-down and clicking on the "NSAID" hyperlink
3. Your patient informed you that they stopped taking fluoxetine. Please remove fluoxetine (Prozac) as a risk factor to show how their risk has changed.	In the patient Risk Profile section, the toggle for "On Selective Serotonin Reuptake Inhibitor" was preconfigured in the on position. The successful action was clicking the toggle off.
4. Assume your patient would like to take a medication then click the button to view medication options.	Navigating to the decision tree questions and clicking on the "Medication" button
5. Your patient has decided to try non-NSAID medication options. Please select acetaminophen (Tylenol) and lidocaine (Lidoderm).	Navigating to the second question of the decision tree and clicking "Other medications" then selecting "acetaminophen (Tylenol) 500mg" and "lidocaine (Lidoderm) 5% patch"
6. Your patient believes NSAIDs help the most with pain but would like to reduce their risk. Please select the oral NSAID option with the least gastrointestinal bleed risk. Then select that a stomach acid reducer is not needed.	Navigating to the second question of the decision tree and clicking on "Oral NSAID" then selecting "celecoxib"
7. Your patient insists on taking medications only once per day. Please select the oral NSAID option with the most risk and add esomeprazole (Nexium).	Navigating to the second question of the decision tree and clicking on "Oral NSAID" and selecting "meloxicam." Then clicking on "Stomach acid reducer" and selecting "esomeprazole."
8. Please place any order in the queue for one of the treatment options.	Navigating through the decision tree and clicking "Accept"

[a]NSAID: nonsteroidal anti-inflammatory drug.

Results

Design

Although DDIs are a novel application for SDM, we did assess several relevant electronic decision aids such as those used for cardiovascular and diabetes management [31]. We compared a variety of relevant decision aids to the SDM steps and the International Patient Decision Aid Standards checklist [26]. Generally, decision aids lacked features to elicit patient values and preferences. Additionally, decision aids varied on the content provided to make treatment decisions including price, effectiveness, and side effects of treatment [32].

Several features from relevant decision aids were adapted to the DDI use case. Features included the icon array, personalized risk, and ability to simulate risk based on patient factors and treatments. Features and functionality evolved through multiple iterations. For example, icon arrays have shown promise in communicating risk to patients and clinicians [33-35]. Figure

2 depicts how the icon array changed over two iterations. Icon array features were based on findings from the literature and expert feedback [36,37]. Once features such as the icon array were acceptable from the design team's perspective, they were adapted and placed into the complete user interface (ie, stable prototype). Figure 3 depicts the stable prototype used for the heuristic evaluation. Overall, 2 major, 13 moderate, and 14 minor issues were identified. One of the two major issues was associated with the number four in Figure 3. The evaluators thought that users may not correctly interpret "Substitute" and "Add" when selecting treatment options. In response, we removed the order entry context, which will avoid a user from referencing the warfarin or NSAID order in process.

Figure 2. Icon array evolution, from left to right, through two iterations. NSAID: nonsteroidal anti-inflammatory drug.

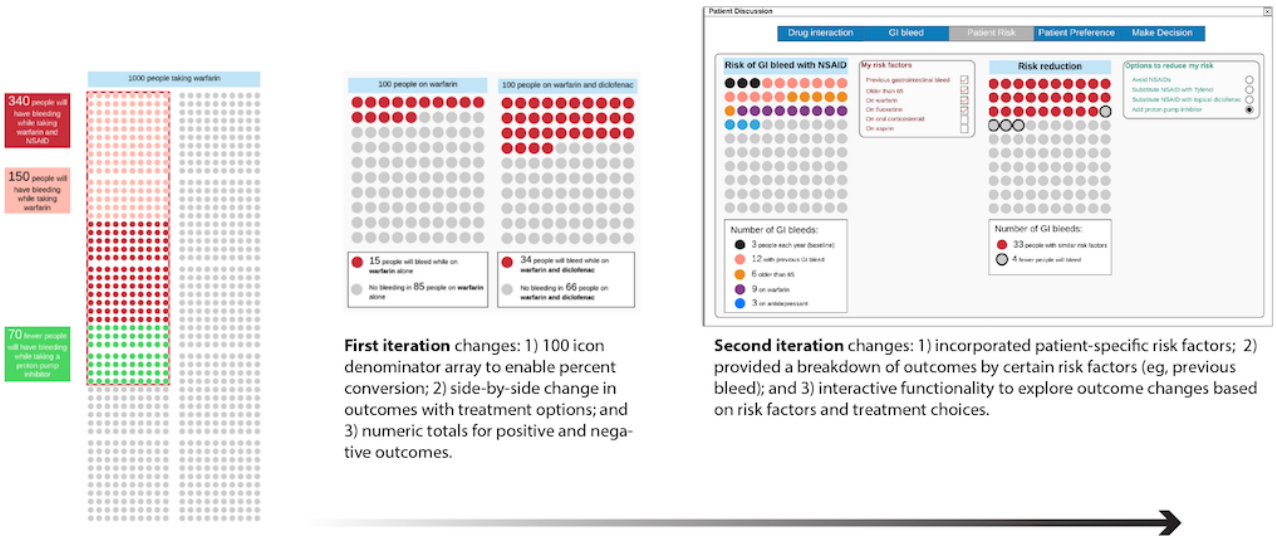
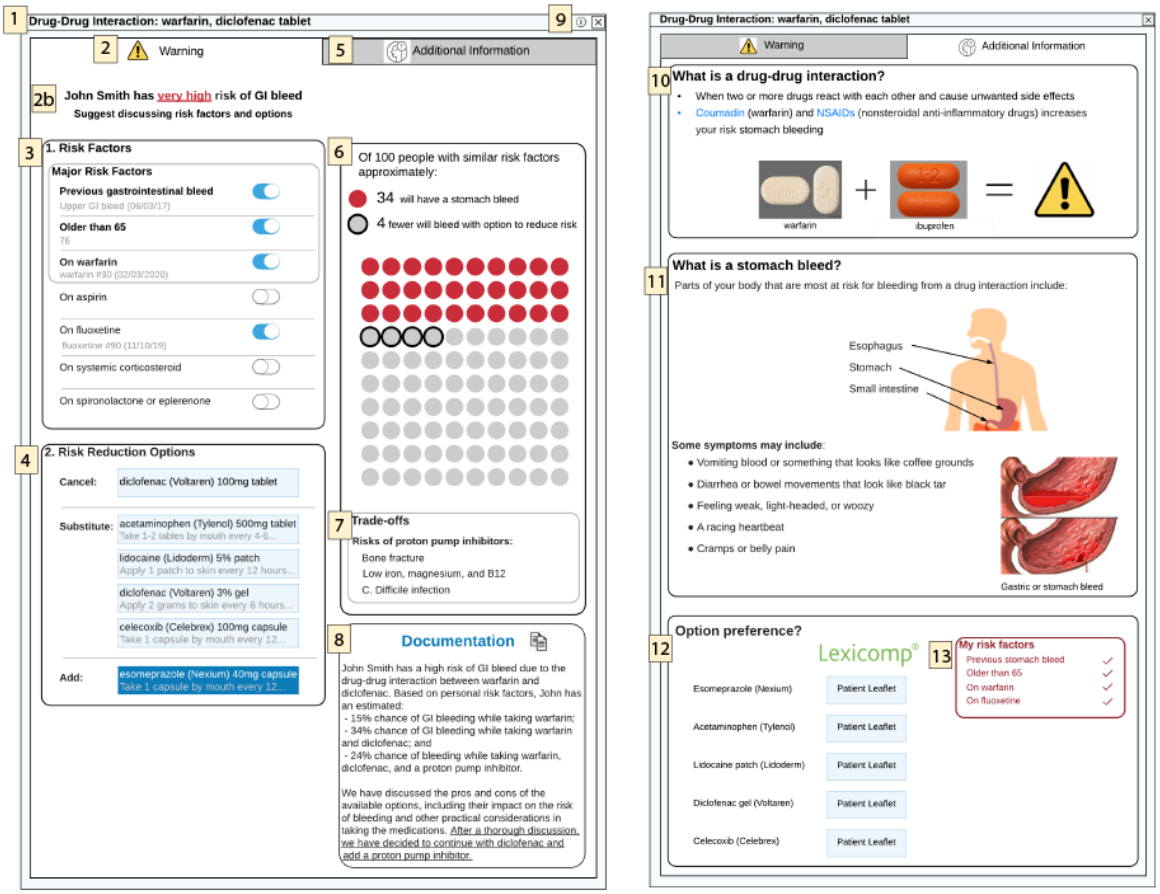
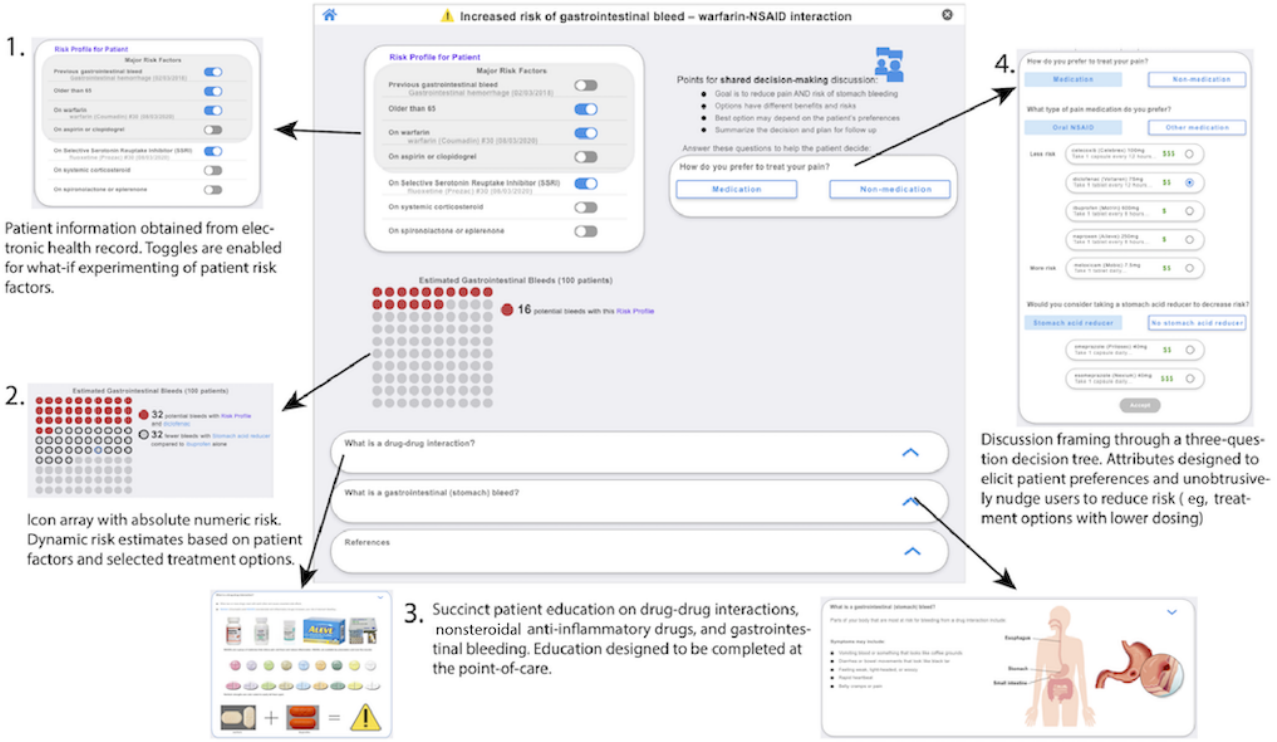


Figure 3. Stable prototype used for heuristic evaluation. Numbers refer to items described in the heuristic report. GI: gastrointestinal.



The DDInteract prototype used for evaluating usability consisted of four sections. The first section in the top left corner of Figure 4 is a patient-specific risk profile. Substantiated risk factors for gastrointestinal bleeding are listed and stratified by risk and supporting evidence. When patient-specific risk factors are pulled from the EHR, the toggle is on (ie, blue). Toggles can be manually changed to account for data not in the EHR or for clinicians to test different scenarios such as adding an antidepressant. The second section in the bottom left is a dynamic risk icon array. The icon array changes with risk factors and treatment options. Absolute numeric risk is provided along the visualization. The third section on the bottom of Figure 4 provides patient education and supporting evidence. Succinct and image-oriented patient education is provided for DDIs and gastrointestinal bleeding. Since risk factors and estimates are evidence based, information is provided on how these aspects were derived. The fourth section is a decision tree with three questions that help structure the conversation and support treatment decision-making. The questions were created and validated by clinicians with the aim to streamline the SDM process in the context of patient care appointments. Question attributes were designed to elicit patient values and preferences associated with bleeding risk and pain treatment. Unobtrusive nudges to reduce risk were used through prepopulated NSAID dosing and default proton pump inhibitor (PPI) selection when an NSAID is chosen. Finally, functionality for generating documentation of the SDM discussion was provided.

Figure 4. Final DDInteract high-fidelity prototype used in the usability study along with feature description. NSAID: nonsteroidal anti-inflammatory drug.



Usability

Patients

All 4 participants were aged 65-85 years and had taken warfarin for more than 5 years. A provider (KM) and a facilitator (author TJR) interviewed each patient for approximately 1 hour. For the high-risk vignette (patient with multiple risk factors), all participants chose a combination of nonmedication treatment (eg, physical therapy) and acetaminophen. For the low-risk scenario (patient with minimal risk factors), most participants chose a short course of celecoxib or ibuprofen, with a PPI.

Although participant knowledge about the warfarin-NSAID DDI varied, all participants appreciated the ability to see DDInteract while the provider discussed risk and treatment options. One participant stated:

> *If I wasn't able to see the [treatment] options, I wouldn't know what to ask*

Furthermore, participants felt empowered to participate in making decisions that aligned with their preferences by referring to the decision aid during discussion:

> *I personally don't like taking medications and want two avoid taking more. It looks like I can try other ways to relieve my pain and I would prefer trying those*

Participants wanted to have access to the decision aid or a printout, outside the encounter, to review what was discussed and decided. One participant stated:

> *I usually forget what we [patient and provider] talk about during the appointment, so I would go to my After Visit Summary to review what we talked about.*

Most participants believed SDM was novel and different from past decision-making experiences with providers:

> *Doctors usually make decisions like these for me.*

Two patients were unaware that the NSAID class included more than ibuprofen. Generally, participants valued SDM and using DDInteract with the provider. Furthermore, participants preferred to avoid additional medications and wanted to reduce the risk of gastrointestinal bleeding as much as possible.

Clinicians

A total of 11 clinicians participated in the usability evaluation (Table 2). Of the 11 participants, 3 stopped the task study after the first task. Of those 3 participants, 2 were pulled to clinical duties. The other participant failed to complete the second task and chose to stop the study rather than skipping the task. Of the 8 participants who completed the study, all were successful on each task. The mean time to complete eight tasks was 144 (SD 74) seconds. Table 3 delineates the task prompt and the mean time for completion. Screen capture was used to determine how participants navigated through the tool. A total of 11 participants completed the usability and satisfaction survey, with an overall mean rating of 4.32 (SD 0.52) of 5. Table 4 delineates mean ratings for each survey item.

Table 2. Participant characteristics for the usability evaluation.

	Participants, n	Specialty (n participants in each group)	Participant years of experience, n				Participant clinical percent effort, n				Self-assessed experience with warfarin from 0 to 100, mean (range)
			<5	6-10	11-15	>16	<21	21-40	61-80	>80	
Physicians	7	Family medicine (4), emergency/critical care (2), hematology (1)	2	2	2	1	0	2	3	2	67 (29-88)
Pharmacists	4	Anticoagulation/ambulatory care (3), general (1)	0	0	2	2	1	0	1	2	93 (87-100)

Table 3. Mean and SD for task time in seconds across usability participants (n=8).

Tasks	Mean time in seconds (SD)
1. Your patient has questions about what a gastrointestinal bleed is. Please navigate to patient education about a gastrointestinal bleeding.	39 (48)
2. With previous patients, you have found it confusing for them to understand the drug class NSAIDs[a]. Please find the picture of multiple NSAIDs to illustrate how not only ibuprofen is an NSAID.	42 (34)
3. Your patient informed you that they stopped taking fluoxetine. Please remove fluoxetine (Prozac) as a risk factor to show how their risk has changed.	2 (1)
4. Assume your patient would like to take a medication then click the button to view medication options.	3 (4)
5. Your patient has decided to try non-NSAID medication options. Please select acetaminophen (Tylenol) and lidocaine (Lidoderm).	30 (48)
6. Your patient believes NSAIDs help the most with pain but would like to reduce their risk. Please select the oral NSAID option with the least gastrointestinal bleed risk. Then select that a stomach acid reducer is not needed.	32 (35)
7. Your patient insists on taking medications only once per day. Please select the oral NSAID option with the most risk and add esomeprazole (Nexium).	13 (4)
8. Please place any order in the queue for one of the treatment options.	15 (10)

[a]NSAID: nonsteroidal anti-inflammatory drug.

Table 4. Clinician usability survey items and responses (n=11). Responses were on a 1 to 5 Likert scale where 1 is strongly disagree and 5 is strongly agree.

Survey items	Mean (SD)
I found the decision tool to be logical.	4.36 (0.67)
I found the decision tool to be efficient.	4.18 (0.75)
The decision tool was effective in the decision-making process.	4.36 (0.67)
The shared decision-making was valuable.	4.27 (0.79)
The decision tool was valuable.	4.36 (0.67)
I thought the decision tool was easy to use.	4.27 (0.65)
I enjoyed the experience.	4.36 (0.81)
I learned something new from this experience.	4.36 (0.67)

Clinician participants provided a variety of comments on the purpose and usefulness of DDInteract after completing the survey. Two participants thought the app would be helpful for patient education:

> *I think half the time they [patients] just think we're [clinicians] being mean by telling them they shouldn't take their NSAIDs. And the visual for how that can be mitigated is great. I actually think that the educational section of the tool would be helpful when we're doing new educations for warfarin/DOACs [direct oral anticoagulants] even if we're not doing a shared decision-making type thing.*

> *Have you considered using this as a tool not just in the clinical setting but in the medical education setting?*

Two participants had questions on where and when the app would be used:

> *Is this an app that will be on the provider's phone or the intention is for the patient to download this app and fill it out themselves? Or will this be a website that is pulled up during an office visit where both parties are present in the room?*

> *I'm not completely clear on the exact clinical situation in which this tool would be used and the point in the workflow in which that would happen.*

One participant thought the tool could be expanded with a general guide on interpreting risk of gastrointestinal bleeding:

> *I know it is individualized, but I like how there are some general guidelines with the HAS-BLED score. It would be nice to have something similar. Also, at what risk is a PPI strongly recommended.*

Finally, participants thought the dynamic risk calculation would be a feature they would return to the tool to use. The ability to toggle between risk factors and treatment options helped to quantify risk and explore different treatment options. Generally, participants believed DDInteract was easy to use and would support SDM.

Prototype Changes

Key changes were made in response to the patient interviews. Changes included expanding nonmedication and non-NSAID treatment options, adding functionality for selecting more than one non-NSAID treatment (eg, physical therapy and acetaminophen), and creating a printable handout and an after visit summary that patients can access outside the encounter. Based on the duration to complete certain tasks and how clinicians navigated through DDInteract, we made feature changes regarding tasks one and two (Table 3). The modifications included enabling the user to see the entire app without scrolling to the drop-down items for patient education and automatically displaying the images.

Discussion

Principal Findings

This study designed and assessed the usability of a tool for SDM with DDIs (Figure 4). Overall, it appears that SDM can be enhanced by using a tool that displays risks of harm and alternatives. The process of designing DDInteract was rigorous, applying user-centered design principles through iterative prototyping (Figure 2). Target users found DDInteract easy to use and believed it could be useful for supporting SDM (Tables 3 and 4). Given that DDI clinical decision support has been traditionally clinician centric, this study may contribute to a major shift in the way certain medication alerts are developed and used. Through the process designing and evaluating DDInteract with clinicians and patients, lessons were learned regarding patient decision-making and their understanding of the warfarin-NSAID DDI. Furthermore, lessons were learned from clinicians about when in the workflow DDI alerts are addressed and implementing SDM in routine patient care.

Not all DDIs are amenable to SDM, and clinicians should use their judgement before opening a discussion about certain DDIs. Although patients appreciated discussing the warfarin-NSAID DDI with the provider, only the low-risk scenario seemed to be relevant. Certain high-risk DDIs should be avoided without SDM. Additionally, although these patients had been on warfarin for multiple years, knowledge about DDIs and treating pain while taking warfarin was limited. This aligns with what others have found on patient knowledge about anticoagulant therapy [38,39]. Knowledge about which medications are NSAIDs and symptoms of bleeding should not be assumed despite experience with warfarin. Regardless of the decision to avoid an NSAID, patient education about the DDI is needed. Finally, patients mentioned that after previous provider encounters it was difficult to recall information about treatment decisions. Consideration for allowing patients to reference the tool after an encounter may help with comprehension and adhering to decisions; however, DDInteract and other similar tools would need to be adapted to and tested with patients to understand decision-making without provider assistance.

Clinicians had questions about when and how DDInteract would be used. Medication prescribing often occurs at the end of an encounter without the patient. If DDInteract was triggered when an NSAID is ordered, the patient might not be available for discussion. Opportunities to use an SDM tool for DDIs earlier in the workflow may be needed. For example, triggering DDInteract for a patient who is having pain or starting warfarin are additional use cases. Triggering on pain is especially relevant for patients on warfarin due to frequent use of over-the-counter NSAIDs. Although DDInteract was designed to mitigate risk associated with DDIs, clinicians requested decision support for other aspects of anticoagulant therapy, such as deciding to start an anticoagulant or which anticoagulant to use. To achieve broad uptake of SDM for DDIs, the scope of DDInteract may need to be expanded to other decision-making and clinician workflow opportunities.

Future Research

The next step is to conduct a formative evaluation of DDInteract to understand how it impacts measures of decision-making, satisfaction, and clinician workflow. To maximize dissemination and enable integration with EHR systems, we have developed an interoperable DDInteract app using emerging clinical decision support standards including Clinical Quality Language, Clinical Decision Support Hooks, and SMART on Fast Healthcare Interoperability Resources (FHIR) [40]. A SMART on FHIR prototype of DDInteract is available on Logica Sandbox, and we have successfully implemented DDInteract in an EHR test environment at the University of Utah. Further research is needed to understand how SDM with DDIs can be integrated with overarching decisions surrounding anticoagulant therapy. Finally, research is needed to explore how decision aids in the EHR can be adapted to clinical workflows to enable SDM in routine patient care.

Conclusion

This study describes the design and usability testing of DDInteract. The findings contribute to knowledge about implementing SDM in routine patient care and expand the use of SDM to DDIs. A multidisciplinary design team collaborated with patients, clinicians, and health information technology experts to design a tool that provides a patient-specific risk calculation, elicits patient preferences, and guides both the patient and clinician to a decision. The rigorous design process resulted in a usable and potentially useful tool. Through the design process and usability testing, key lessons were learned from the patient and clinician perspectives. The next step is to evaluate the utility of DDInteract in a clinical setting, and if successful, to implement it across EHRs using interoperable technology.

Conflicts of Interest

This study was funded by the Agency for Healthcare Research and Quality (U18 PA-18-792 and R18HS026198). Outside of the described work, KK reports honoraria, consulting, sponsored research, licensing, or codevelopment in the past 3 years with McKesson InterQual, Hitachi, Pfizer, Premier, Klesis Healthcare, RTI International, Mayo Clinic, the University of Washington, the University of California at San Francisco, MD Aware, and the US Office of the National Coordinator for Health IT (via ESAC and Security Risk Solutions) in the area of health information technology. KK was also an unpaid board member of the nonprofit Health Level Seven International health information technology standard development organization, he is an unpaid member of the US Health Information Technology Advisory Committee, and he has helped develop a number of health information technology tools outside of the described work, which may be commercialized to enable wider impact. None of these relationships have direct relevance to the paper but are reported in the interest of full disclosure. Other authors have no competing interests to declare.

Multimedia Appendix 1
Clinic vignettes used for patient interviews.
[DOCX File , 14 KB - humanfactors_v8i4e28618_app1.docx]

References

1. Magro L, Moretti U, Leone R. Epidemiology and characteristics of adverse drug reactions caused by drug-drug interactions. Expert Opin Drug Saf 2012 Jan;11(1):83-94. [doi: 10.1517/14740338.2012.631910] [Medline: 22022824]
2. Hines LE, Murphy JE. Potentially harmful drug-drug interactions in the elderly: a review. Am J Geriatr Pharmacother 2011 Dec;9(6):364-377. [doi: 10.1016/j.amjopharm.2011.10.004] [Medline: 22078863]
3. Juurlink DN, Mamdani M, Kopp A, Laupacis A, Redelmeier DA. Drug-drug interactions among elderly patients hospitalized for drug toxicity. JAMA 2003 Apr 02;289(13):1652-1658. [doi: 10.1001/jama.289.13.1652] [Medline: 12672733]
4. Qato DM, Alexander GC, Conti RM, Johnson M, Schumm P, Lindau ST. Use of prescription and over-the-counter medications and dietary supplements among older adults in the United States. JAMA 2008 Dec 24;300(24):2867-2878 [FREE Full text] [doi: 10.1001/jama.2008.892] [Medline: 19109115]
5. Hastings S, Schmader K, Sloane R, Weinberger M, Pieper CF, Goldberg KC, et al. Quality of pharmacotherapy and outcomes for older veterans discharged from the emergency department. J Am Geriatr Soc 2008 May;56(5):875-880. [doi: 10.1111/j.1532-5415.2008.01648.x] [Medline: 18341519]
6. Hanlon JT, Perera S, Newman AB, Thorpe JM, Donohue JM, Simonsick EM, Health ABC Study. Potential drug-drug and drug-disease interactions in well-functioning community-dwelling older adults. J Clin Pharm Ther 2017 Apr;42(2):228-233 [FREE Full text] [doi: 10.1111/jcpt.12502] [Medline: 28111765]
7. Delaney JA, Opatrny L, Brophy JM, Suissa S. Drug drug interactions between antithrombotic medications and the risk of gastrointestinal bleeding. CMAJ 2007 Aug 14;177(4):347-351 [FREE Full text] [doi: 10.1503/cmaj.070186] [Medline: 17698822]
8. Malone D, Hutchins D, Haupert H, Hansten P, Duncan B, Van Bergen RC, et al. Assessment of potential drug-drug interactions with a prescription claims database. Am J Health Syst Pharm 2005 Oct 01;62(19):1983-1991. [doi: 10.2146/ajhp040567] [Medline: 16174833]
9. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006;13(2):138-147 [FREE Full text] [doi: 10.1197/jamia.M1809] [Medline: 16357358]

10. Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS. Physicians' decisions to override computerized drug alerts in primary care. Arch Intern Med 2003 Nov 24;163(21):2625-2631. [doi: 10.1001/archinte.163.21.2625] [Medline: 14638563]
11. Bryant A, Fletcher G, Payne T. Drug interaction alert override rates in the Meaningful Use era: no evidence of progress. Appl Clin Inform 2014;5(3):802-813 [FREE Full text] [doi: 10.4338/ACI-2013-12-RA-0103] [Medline: 25298818]
12. Heringa M, Floor-Schreudering A, Wouters H, De Smet PAGM, Bouvy ML. Preferences of patients and pharmacists with regard to the management of drug-drug interactions: a choice-based conjoint analysis. Drug Saf 2018 Feb;41(2):179-189 [FREE Full text] [doi: 10.1007/s40264-017-0601-7] [Medline: 28965265]
13. Dolan JG, Veazie PJ, Russ AJ. Development and initial evaluation of a treatment decision dashboard. BMC Med Inform Decis Mak 2013 Apr 21;13:51 [FREE Full text] [doi: 10.1186/1472-6947-13-51] [Medline: 23601912]
14. Tiedje K, Shippee ND, Johnson AM, Flynn PM, Finnie DM, Liesinger JT, et al. 'They leave at least believing they had a part in the discussion': understanding decision aid use and patient-clinician decision-making through qualitative research. Patient Educ Couns 2013 Oct;93(1):86-94 [FREE Full text] [doi: 10.1016/j.pec.2013.03.013] [Medline: 23598292]
15. Sandman L, Munthe C. Shared decision making, paternalism and patient choice. Health Care Anal 2010 Mar;18(1):60-84. [doi: 10.1007/s10728-008-0108-6] [Medline: 19184444]
16. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
17. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997 Mar;44(5):681-692. [doi: 10.1016/s0277-9536(96)00221-3] [Medline: 9032835]
18. Légaré F, Adekpedjou R, Stacey D, Turcotte S, Kryworuchko J, Graham ID, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2018 Jul 19;7:CD006732 [FREE Full text] [doi: 10.1002/14651858.CD006732.pub4] [Medline: 30025154]
19. Wright M, Dunbar S, Macpherson B, Moretti E, Del Fiol G, Bolte J, et al. Toward designing information display to support critical care. A qualitative contextual evaluation and visioning effort. Appl Clin Inform 2016 Oct 05;7(4):912-929. [doi: 10.4338/ACI-2016-03-RA-0033] [Medline: 27704138]
20. Reese TJ, Kawamoto K, Del Fiol G, Weir C, Tonna J, Segall N, et al. Approaching the design of an information display to support critical care. 2017 Aug 23 Presented at: 2017 IEEE International Conference on Healthcare Informatics; August 23-26, 2017; Park City, UT. [doi: 10.1109/ichi.2017.64]
21. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. Encyclopedia of Human-Computer Interaction. Thousand Oaks, CA: Sage Publications; 2004:445-456.
22. Mao J, Vredenburg K, Smith PW, Carey T. The state of user-centered design practice. Commun ACM 2005 Mar 01;48(3):105-109. [doi: 10.1145/1047671.1047677]
23. Nelson S, Del Fiol G, Hanseler H, Crouch B, Cummins M. Software prototyping: a case report of refining user requirements for a health information exchange dashboard. Appl Clin Inform 2016;7(1):22-32 [FREE Full text] [doi: 10.4338/ACI-2015-07-CR-0091] [Medline: 27081404]
24. Heaton N. What's wrong with the user interface: how rapid prototyping can help. 1992 Nov 11 Presented at: IEE Colloquium on Software Prototyping and Evolutionary Development; November 11, 2002; London, UK p. 1-7.
25. The SHARE Approach—essential steps of shared decisionmaking: quick reference guide. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/health-literacy/professional-training/shared-decision/tools/resource-1.html [accessed 2020-11-15]
26. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, International Patient Decision Aids Standards (IPDAS) Collaboration. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006 Aug 26;333(7565):417 [FREE Full text] [doi: 10.1136/bmj.38926.629329.AE] [Medline: 16908462]
27. Nielsen J. Usability inspection methods. In: Conference Companion on Human Factors in Computing Systems. 1994 Apr 28 Presented at: CHI '94; April 24-28, 1994; Boston, MA p. 413-414. [doi: 10.1145/259963.260531]
28. Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1990 Mar 01 Presented at: CHI '90; April 1-5, 1990; Seattle, WA p. 249-256. [doi: 10.1145/97243.97281]
29. Brooke J. SUS: a 'Quick and Dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. London: CRC Press; 1996.
30. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
31. Mayo Clinic Shared Decision Making National Resource Center. care that fits. URL: https://carethatfits.org/shared-decision-making/ [accessed 2020-11-14]
32. O'Connor AM, Bennett C, Stacey D, Barry MJ, Col NF, Eden KB, et al. Do patient decision aids meet effectiveness criteria of the international patient decision aid standards collaboration? A systematic review and meta-analysis. Med Decis Making 2007;27(5):554-574. [doi: 10.1177/0272989X07307319] [Medline: 17873255]

33. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. Health Psychol 2009 Mar;28(2):210-216. [doi: 10.1037/a0014474] [Medline: 19290713]
34. Sheridan SL, Halpern DJ, Viera AJ, Berkman ND, Donahue KE, Crotty K. Interventions for individuals with low health literacy: a systematic review. J Health Commun 2011;16 Suppl 3:30-54. [doi: 10.1080/10810730.2011.604391] [Medline: 21951242]
35. Garcia-Retamero R, Galesic M, Gigerenzer G. Do icon arrays help reduce denominator neglect? Med Decis Making 2010;30(6):672-684. [doi: 10.1177/0272989X10369000] [Medline: 20484088]
36. Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PKJ, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak 2013;13 Suppl 2:S7 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S7] [Medline: 24625237]
37. Spiegelhalter D, Pearson M, Short I. Visualizing uncertainty about the future. Science 2011 Sep 09;333(6048):1393-1400. [doi: 10.1126/science.1191181] [Medline: 21903802]
38. Taylor FC, Ramsay ME, Tan G, Gabbay J, Cohen H. Evaluation of patients' knowledge about anticoagulant treatment. Qual Health Care 1994 Jun;3(2):79-85 [FREE Full text] [doi: 10.1136/qshc.3.2.79] [Medline: 10137589]
39. Tang EOY, Lai CS, Lee KK, Wong RS, Cheng G, Chan TY. Relationship between patients' warfarin knowledge and anticoagulation control. Ann Pharmacother 2003 Jan;37(1):34-39. [doi: 10.1345/aph.1A198] [Medline: 12503930]
40. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016 Sep;23(5):899-908 [FREE Full text] [doi: 10.1093/jamia/ocv189] [Medline: 26911829]

Abbreviations

DDI: drug-drug interaction
DOAC: direct oral anticoagulant
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
NSAID: nonsteroidal anti-inflammatory drug
PPI: proton pump inhibitor
SDM: shared decision-making

Edited by A Kushniruk; submitted 08.03.21; peer-reviewed by K Trinkley, D Chrimes, S Pelayo; comments to author 04.07.21; revised version received 19.07.21; accepted 25.07.21; published 26.10.21.

Please cite as:
Reese TJ, Del Fiol G, Morgan K, Hurwitz JT, Kawamoto K, Gomez-Lumbreras A, Brown ML, Thiess H, Vazquez SR, Nelson SD, Boyce R, Malone D
A Shared Decision-making Tool for Drug Interactions Between Warfarin and Nonsteroidal Anti-inflammatory Drugs: Design and Usability Study
JMIR Hum Factors 2021;8(4):e28618
URL: https://humanfactors.jmir.org/2021/4/e28618
doi:10.2196/28618
PMID:34698649

Original Paper

Medical Specialists' Perspectives on the Influence of Electronic Medical Record Use on the Quality of Hospital Care: Semistructured Interview Study

Rube van Poelgeest[1], MSc; Augustinus Schrijvers[1], Prof Dr; Albert Boonstra[2], Prof Dr; Kit Roes[3], Prof Dr

[1]Julius Center, University Medical Center, University of Utrecht, Utrecht, Netherlands
[2]Faculty of Economics and Business, University of Groningen, Groningen, Netherlands
[3]Radboudumc, University of Nijmegen, Nijmegen, Netherlands

Corresponding Author:
Rube van Poelgeest, MSc
Julius Center
University Medical Center
University of Utrecht
Universiteitsweg 100
Utrecht, 3584 CG
Netherlands
Phone: 31 620139545
Email: rube.van.poelgeest@planet.nl

Abstract

Background: Numerous publications show that electronic medical records (EMRs) may make an important contribution to increasing the quality of care. There are indications that particularly the medical specialist plays an important role in the use of EMRs in hospitals.

Objective: The aim of this study was to examine how, and by which aspects, the relationship between EMR use and the quality of care in hospitals is influenced according to medical specialists.

Methods: To answer this question, a qualitative study was conducted in the period of August-October 2018. Semistructured interviews of around 90 min were conducted with 11 medical specialists from 11 different Dutch hospitals. For analysis of the answers, we used a previously published taxonomy of factors that can influence the use of EMRs.

Results: The professional experience of the participating medical specialists varied between 5 and 27 years. Using the previously published taxonomy, these medical specialists considered technical barriers the most significant for EMR use. The suboptimal change processes surrounding implementation were also perceived as a major barrier. A final major problem is related to the categories “social” (their relationships with the patients and fellow care providers), “psychological” (based on their personal issues, knowledge, and perceptions), and “time” (the time required to select, implement, and learn how to use EMR systems and subsequently enter data into the system). However, the medical specialists also identified potential technical facilitators, particularly in the assured availability of information to all health care professionals involved in the care of a patient. They see promise in using EMRs for medical decision support to improve the quality of care but consider these possibilities currently lacking.

Conclusions: The 11 medical specialists shared positive experiences with EMR use when comparing it to formerly used paper records. The fact that involved health care professionals can access patient data at any time they need is considered important. However, in practice, potential quality improvement lags as long as decision support cannot be applied because of the lack of a fully coded patient record.

(JMIR Hum Factors 2021;8(4):e27671) doi:10.2196/27671

KEYWORDS

electronic medical record (emr); hospitals; quality; health care; medical specialist

Introduction

In modern-day hospitals, IT is present in many forms. Among these are information systems, networks, databases, and websites. An electronic medical record (EMR) comprehensively includes all information to support medical diagnosis and treatment within the same institution or health system. EMR use generally alludes to the transition of information to a digital form, that is, a form that can be used by electronic devices, such as computers. Various authors agree that EMRs can make an important contribution to increasing the quality of care [1,2]. However, how are EMR use and the quality of medical care related in a hospital context?

In previous studies, the two authors (RvP and KR) of this paper attempted to establish links between the extent of EMR use and the quality of medical care [3-5]. In those previous quantitative studies, they used a specially developed tool to measure the degree of EMR use in Dutch hospitals. This 8-stage (0-7) measurement tool, the so-called EMR Adoption Model (EMRAM) from the Healthcare Information and Management Systems Society (HIMSS) measures the adoption and use of EMR functions. EMRAM incorporates algorithms to score hospitals relative to their EMR capabilities and aims to encourage hospitals to use EMRs at a higher stage [6]. The HIMSS is a US not-for-profit organization dedicated to improving health care in its quality, safety, cost-effectiveness, and access through the best use of IT and management systems. It was founded in 1961 as the Hospital Management Systems Society. It is now headquartered in Chicago, IL, USA. The society has more than 80,000 individuals, 480 provider organizations, 470 nonprofit partners, and 650 health services organizations (as of December 2019). The HIMSS definition of an environment with a complete EMR (stage 7) is “an environment that is composed of the clinical data repository, clinical decision support, controlled medical vocabulary, order entry, computerized practitioner order entry, and clinical and physician documentation applications” [7]. Ultimately, the model should lead to the use of EMR systems so that the hospital no longer uses paper charts. The findings of the quantitative analysis in the previous studies [3-5] show that Dutch hospitals in 2014 particularly struggled with electronic nursing documentation. In 2012-2014, 37.5% of Dutch hospitals were unable to upload this information to the EMRs. Once this challenge is met, the next challenge for Dutch hospitals will be to equip the EMRs with closed loop medication administration (CLMA) and an advanced clinical decision support system (CDSS). The CLMA is a fully electronic medication management process in which all relevant information is seamlessly documented. All the steps in the medication cycle (ordering, verifying, preparing, and administering) are supported electronically with decision support, where relevant. A CDSS is an application that analyzes data to help health care providers make decisions and improve patient care. A CDSS focuses on using knowledge management to obtain clinical advice based on multiple factors of patient-related data. It enables integrated workflows, provides assistance at the time of care, and offers care plan recommendations.

A 2015 study [4] tried to find a correlation between the EMRAM score and Elsevier performance indicators. This yearly Elsevier publication is a Dutch nationwide publication of quality indicators for hospitals. No statistically significant correlations were found.

In the 2017 study [5], a positive association between the use of EMRs and patient quality outcomes was found for the length of stay (LOS, the duration of a single episode of hospitalization) for patients with colorectal cancer in Dutch hospitals, as measured by the Dutch Surgical Colorectal Audit (DSCA).

In a third study (2018, not yet published, available from the first author [RvP]), we did not find a significant relationship between the EMRAM score and the number of patients with adverse events (AEs), preventable AEs, AEs caused by medication, number of re-admissions (RAs), and the LOS, as measured in the NIVEL study [8]. Our research team did not understand why a better EMRAM score does not lead to a better quality of care. We believed that two intervening aspects play a role: the implementation process itself and the role of medical specialists. We believed so because of several publications on both aspects.

The first of these is a paper written by Adler-Milstein et al [9], which emphasizes the importance of the implementation process of more mature IT systems for reaching higher quality. Recent studies suggest that unsuccessful implementation of EMR systems could be due to poorly designed EMR systems, poor use of EMRs by clinicians, or social organizational aspects, such as goal conflicts, lack of time, or lack of support from colleagues [10]. The second factor is the role of the medical specialist [11,12]. Previous studies show this is an important factor in the adoption and use [6] of EMR systems in hospitals. Medical specialists are a main frontline group of users of EMR systems. In addition, whether they support and effectively use EMR systems greatly influences other user groups in the medical institution, such as nurses, pharmacists, and administrative staff. To optimize EMR use, it is therefore essential to understand what physicians perceive to be key aspects that either support or hinder the use of EMR systems, which can positively impact medical treatment and care. To substantiate our ideas about not finding a relationship between the EMRAM score and the quality of care, we undertook this study with the following open research question:

> *Which positive or negative aspects influence the relationship between EMR use and the quality of medical care according to medical specialists?*

In this paper, the term “EMR” can concern the data themselves, the accompanying procedures, or a fundamental change in method (so-called digital transformation).

Methods

Design and Methodology

To answer the research question, a qualitative research study was performed. A qualitative design was considered appropriate for this question, as the primary objective was to explore more in-depth perceptions of factors and processes related to a more complex system, including social and technical components [13].

The development of EMR use in 73 Dutch hospitals in the period of 2012-2015 was measured using the EMRAM score [14]. The hospitals that were measured twice in the research period and did not work with nursing documentation in EMRs or with the CLMA and advanced CDSS were asked to participate in a follow-up study (Table 1). Two hospitals achieved a higher stage on the EMRAM score, eight hospitals stayed on the same level, and one did not respond (Table 1).

Table 1. Summary of participating hospitals based on the EMRAM[a] score.

Participant number	Type of hospital	Nursing documentation and CLMA[b]/advanced CDSS[c] in EMRs[d] in 2012-2014	Nursing documentation and CLMA/advanced CDSS in EMRs in 2018
1	UMC[e]	No	No
2	UMC	No	Yes
3	Teaching hospital	No	No
4	Teaching hospital	No	No
5	Teaching hospital	No	N/A[f]
6	Teaching hospital	No	No
7	Local hospital	No	No
8	Local hospital	No	No
9	Local hospital	No	No
10	Local hospital	No	No
11	Local hospital	No	Yes

[a]EMRAM: Electronic Medical Record Adoption Model.

[b]CLMA: closed loop medication administration.

[c]CDSS: clinical decision support system.

[d]EMR: electronic medical record.

[e]UMC: University Medical Centre.

[f]N/A: not available.

The hospitals were approached through the chairs of their medical staff and were asked to nominate a medical specialist each to participate in this study. To be eligible, the medical specialists were required to have worked for more than 5 years in the hospital in question. Overall, the selection of participants was geared toward a balanced mix of different specialties (surgical, nonsurgical, small specialty).

In the period of August-October 2018, a semistructured interview of about 90 min was conducted with each medical specialist selected. The abovementioned research question was at the core of the interview. An item list (available from the first author [RvP]) was used by the interviewer to help the participants focus on relevant experiences in case the conversation halted. We only asked questions about aspects the medical specialists were personally dealing with; we were not interested in second-hand accounts.

For analysis of the answers, we used the classification of aspects that can influence the implementation of EMR systems based on the taxonomy of Boonstra et al [15]. This systematic literature review was carried out to identify all the barriers that result in physicians showing resistance toward EMR systems. Table 2 shows the highlights of this taxonomy model.

We used this taxonomy for the same aspects in a neutral connotation, as the same taxonomy can be followed when categorizing aspects (quotes) as facilitators of EMR use. All authors participated in the allocation of quotes from the interviews to a category of the described taxonomy, with initial allocation by the first author (RvP) and validation by the second and third authors (AJPS and AB). The results were recorded in a Microsoft Excel file, which is available upon request from the first author (RvP). Based on this classification and the primary interview recordings, the authors reached a consensus about the best way to allocate quotes to the categories of the taxonomy. All authors agreed with the final allocation of quotes.

Table 2. Summary of categories.

Quote category	Description
Technical	The technical aspects of the systems, the technical capabilities of the physicians and the suppliers
Psychological	Concerns regarding the use of EMRs[a] that are based on the medical specialists' personal issues, knowledge, and perceptions
Social	Relationships with the patients and fellow care providers but also with suppliers, insurers, and politicians
Time	Time required to select, implement, and learn how to use EMR systems and subsequently enter data into the system
Finance	Financial issues, including those related to monetary issues in implementing EMR systems
Legal	Privacy or security concerns regarding patients' medical information
Organization	Organizational characteristics, such as size and type of individual practices
Change process	The influence of the organizational culture, incentives, community-level participation, and leadership

[a]EMR: electronic medical record.

Ethics Approval and Consent to Participate

As this study did not involve research on human subjects, no medical ethical committee approval was required under Dutch law. Neither the Dutch Medical Research Involving Human Subjects Act (Wet Medisch-Wetenschappelijk Onderzoek met Mensen [WMO]) nor the university required ethics approval for the type of work conducted in this research. All participants orally and voluntarily agreed to participate in this study. They allowed us to use the data they provided, including quotes, under the condition of confidentiality. All participants agreed with the final report of their interviews. All participating hospitals, medical specialists, and used quotes in the manuscript were anonymized. No written permission was needed in this case.

Results

Participating Hospitals and Medical Specialists

In all, 11 hospitals (Table 3) agreed to participate in this qualitative study, while 3 hospitals (1 regional and 2 teaching hospitals) were unwilling or unable to participate. Each participating hospital nominated 1 medical specialist to be interviewed, representing 10 different specialties and between 5 and 27 years of experience in their current hospital. Six of them were (former) chairs of medical staff.

Table 3. Summary of participating hospitals and medical specialists.

Type of hospital	Number of hospitals (n)	Number of specialists (n)	Type of specialist (n)	Age (years)	Gender (n)	Number of years of experience in hospital
UMC[a]	2	732-1050	Internist (1) Anesthetist (2)	43-57	Female (1) Male (1)	10-13
Teaching hospital	4	240-377	Rheumatologist (3) Radiologist (4) Internist (5) Surgeon (6)	49-62	Male (2) Female (2)	11-23
Local hospital	5	70-187	Pediatric neurologist (7) Vascular surgeon (8) Gynecologist (9) Cardiologist (10) Pharmacist (11)	42-59	Male (5)	5-27

[a]UMC: University Medical Centre.

In total, the participants made 160 observations regarding aspects that influence the relationship between the extent of EMR use and the quality of care: 122 observations were characterized as *barriers* and 38 as *facilitators*. First, we will discuss the technical aspects that are mentioned most often. Next, we will discuss the other aspects of EMRs that influence the quality of care. Not every aspect of the taxonomy was used, because some aspects were not mentioned during the interviews. The legal aspect, related to information safety, was not mentioned in any of the interviews. The organizational aspect (type and size of the hospital) could not be addressed in the analysis because it was not part of the design of the study and was out of scope for individual participants. As explained in the Methods section, the change process aspect is of particular interest from a systems perspective, and it will be treated last to reflect on possible future developments.

The participating hospitals included two academic hospitals, four teaching hospitals, and five local hospitals (see also Tables 1 and 3). However, based on hospital type, no difference was

found between participants' observations. The availability of nursing documentation and of the CLMA and advanced CDSS in the hospital concerned did not lead to differences (see also Tables 1 and 3) in the experiences of the medical specialists.

Technical Aspects

During the interviews, technical aspects were the category all medical specialists chose to talk about first. It therefore seems like medical specialists consider technical aspects the most important factor influencing EMR use in hospitals. To provide more insight, the related tables include a subdivision of technical aspects, followed by quotes from participants to illustrate what the more abstract terms of the model mean. A complete list of all quotes is available from the first author (RvP).

Customizability

Customizability is the ability of the technology system to adapt to specific needs of the user. Within the technical aspects category, customizability (Table 4) was mentioned most often, more often as a barrier than as a facilitator. The medical specialists compared EMR systems in the hospital with intelligent systems that can be used at home to buy a book or book a trip. It seems as though an administrative system has simply been converted to a medical system.

Table 4. Customizability: illustrative quotes from participants.

Quote type	Quote
Barrier	• "Not intuitive. Terrible user interface. Unpleasant system, it clearly hasn't been primarily designed for doctors and paramedics. An originally administrative system that has been reshaped into a medical system." (Participant 2) • "It is digital, but that about says it all. Leaves much to be desired." (Participant 9) • "We can see the added value, but these systems are shoddy. Not intuitive." (Participant 8) • "Preoperative polio. Supplementary lab research takes 1-2 days. If you want to change policy based on the results, the EMR system shows that this is impossible because the patient has not been hospitalized but is not present at the outpatient clinic either." (Participant 2)
Facilitator	• "Innumerable positive points; accessible everywhere, even at home. No more illegible notes." (Participant 10) • "Back in the day, the paper records often got lost. Lab results are available more quickly now, and the medical specialist can quickly see the daily reports of the nurses," (Participant 7)

Interconnectivity/Standardization

EMR hardware and software can be used straight out of the box, but they have to be interconnected with other devices that complement the EMR system. The exchange of dossiers is often not possible due to lack of standardization, and files are often split up between different specialties because that is how the medical practice is organized in the hospital (Table 5). Getting an integral view of a patient's situation is therefore difficult but especially important with multimorbid patients (an evergrowing group). General data, such as blood pressure, smoking, and alcohol use, are often contradictory and recorded more than once. Sometimes, multiple systems have to be simultaneously used during treatment because files are not linked—a situation that medical specialists consider potentially dangerous.

Table 5. Interconnectivity: illustrative quotes from participants.

Quote type	Quote
Barrier	• "Many separate systems are high risk because they are not linked, for example. when transferring files." (Participant 11) • "Gynecologists work with 4 systems (safety risk) because systems are not interlinked." (Participant 9) • "EMR[a] now strongly split into specialties." (Participant 4)
Facilitator	• "Good-quality photos can be easily obtained." (Participant 7) • "Back in the day, there was no background information available if the GP's[b] notification read 'diarrhea'; now there is." (Participant 5)

[a]EMR: electronic medical records

[b]GP: general physician.

Limitations of the System

According to the participants, the IT system promises a great deal but offers little more than the old paper situation. Participants particularly point to the promised additional intelligence that is either absent from the system or present in a limited sense (Table 6). The system could offer, for example, so-called evidence-based advice based on the individual and combined patient data available in the system [16]. An oft-heard theme is also the lack of analytical tools to analyze the available data and to anticipate developments in the health of patients in the hospital.

Table 6. Limitations of the IT system: illustrative quotes from participants.

Quote type	Quote
Obsolescence: the IT system reaching its limit, becoming obsolete, and no longer remaining useful	
Barrier	• "Actually, no added value, no decision support." (Participant 5) • "The hospital world can still learn a lot from, for example, the travel industry. It is madness that you can book a holiday in Thailand within an hour, but that scheduling an operation for a patient with a serious condition causes so many problems." (Participant 2) • About the use of analytic tools: "Executing the analysis was very time consuming. Analytics have to be carried out by an IT specialist. This makes it a hopeless affair. These tools should be included in the EMR[a]." (Participant 9)
Facilitator	• "There is a little bit of decision support for medication (prescriptions)." (Participant 5)
Complexity: EMRs resulting in physicians having to allocate time and effort if they are to master complexity	
Barrier	• "Reporting of transactions [is] very complicated." (Participant 1)
Facilitator	• N/A[b]
Reliability: the dependability of the IT system	
Barrier	• "In [the] case of failure of systems at polyclinic, nothing is available anymore." (Participant 7)
Facilitator	• "Simple but works well. Very few malfunctions. Much better than paper." (Participant 9)

[a]EMR: electronic medical record.

[b]N/A: not available.

Other Aspects of EMRs Influencing the Quality of Care

In Table 7, other aspects of EMRs are summarized. These aspects were mentioned less often by the medical specialists but are also important influencing factors. These observations are generally consistent with the findings published elsewhere [17,18]. One thing that stands out is that medical specialists sometimes miss informal contacts of meetings with colleagues that were previously necessary due to the lack of a common digital file. Equally striking is that the financial aspect was hardly mentioned. The latter contrasts with findings in other publications [16].

Table 7. Other aspects of EMRs[a]: illustrative quotes from participants.

Quote type	Quote
Computer skills of the physician or staff: technical knowledge and skills to deal with EMRs	
Barrier	• "Doctors are not IT savvy, [for] example, [a] radiologist who wants to look at photos at home but uses a home PC[b] that has not been updated and therefore does not work properly." (Participant 8)
Facilitator	• N/A[c]
Training and support: associated with the EMR system	

Barrier	• "The system may not be used properly by medical specialist[s]:
	  • Too little knowledge of the system
	  • Possibilities not known
	  • Defensive medicine: hedge behavior." (Participant 4)

Facilitator	• N/A
Psychological: personal issues, knowledge, and perceptions	
Barrier	• "Check lists: system steers behavior. Against check lists: action is carried out anyway because I have prescribed it. Medication verification is standard procedure, so why check?" (Participant 5) • "There is a contrast between old and young specialists. I think the older ones accept a limited system more easily; their demands are less high." (Participant 7)
Facilitator	• "Enforces a certain treatment, and that is positive." (Participant 10)
Social: relationships with patients and fellow care providers but also with suppliers, insurers, and politicians	
Barrier	• "Medical specialists clearly have ideas about each other. A lot of contradictions. Hard to get on the same page." (Participant 2) • "Back in the day, photos sometimes disappeared (dangerous), but medical specialists came to radiology because there was only one photo; this meant people knew each other, radiology was the center, people walked in, it used to run more smoothly. Now there is multidisciplinary consultation, but people don't know each other anymore." (Participant 4)
Facilitator	• "Member of medical staff (gynecologist) mans a so-called 'wailing wall'." (Participant 8)

[a]EMR: electronic medical record.

[b]PC: personal computer.

[c]N/A: not available.

The Change Process Aspect

In Table 8, illustrative quotes for the change process aspect are summarized. According to the medical specialists, several preconditions for success must be met before the successful introduction of EMRs in their hospitals. There is some doubt, for example, as to whether the supplier of EMRs is willing to create links to other parts of the IT system. However, this runs counter to market forces. Moreover, the participants mentioned that governmental institutions often also still require medical specialists to use paper. A central theme for almost all interviewed medical specialists is the coded or noncoded recording of obtained information. They generally realize that encoding a medical record is a prerequisite for getting help from the EMR system based on so-called evidence-based material. Several hospitals initially started out with the recording of this information by medical specialists but later abolished this system because the medical specialists refused to work with it.

Table 8. The change process: illustrative quotes from participants.

The change process	Quote
Support from organizational culture	• "There is too little attention for resistance in medical specialists due to, for example, time pressure." (Participant 8) • "Before, medical specialists were individual, had their own working methods. By now, a technological revolution has taken place (paper records are now electronic records). But people do not want to change (95%). They have to get out of their comfort zone. You have to invest in that. Now, medical specialists' approach EMR[a] as if it were paper." (Participant 10) • "The problem is that hospitals are not IT minded. Hospitals are not flexible." (Participant 10)
Leadership	• "On its own, the market will not provide properly functioning IT systems for hospitals." (Participant 2) • "Cytostatic control by pharmacies should be done via inspection on paper." (Participant 8) • "Participant sees movement from specialism based (departments) towards disease related. For example, a department of bowel cancer with [an] internist, [an] MDL[b], [an] oncologist, and [a] radiologist. This has an impact on the way digitization is organized." (Participant 4) • "(EMR supplier mentioned) is monopolist. Does not listen to customer." (Participant 4)
Incentives	• "On its own, the market will not provide properly functioning IT systems for hospitals, for example, the market does not benefit from the exchangeability of data. Market forces therefore do not lead to the solution." (Participant 2) • "No 'reward' for 'good' use." (Participant 8)
Participation	• "Conclusion: Letting medical specialists do coding work is undesirable, but it is necessary to enable systems to 'offer help' on a more advanced level. This process should be structured differently by giving supporting staff a role in it." (Participant 8)

[a]EMR: electronic medical record.

[b]MDL: *maag-darm-lever* in Dutch, meaning gastroenterologist.

Discussion

Preliminary Findings

To answer the research question "Which positive or negative aspects influence the relationship between EMR use and the quality of medical care according to medical specialists?", a qualitative research study was performed.

The overall picture of the relationship between the extent of EMR use and the quality of medical care according to the participants shows that medical specialists prefer digital records over old paper ones. However, at the same time, the participants consider the technical systems old-fashioned compared to the systems they can access at home to book a trip or buy a book. The inability of all those involved (professional groups, boards, suppliers, politicians) to improve this situation is described openly by some participants. By and large, the participants do see the potential, but a better way to record coded information still needs to be found. The lack of interconnection between the different EMRs, for example, per hospital department (eg, internal medicine and cardiology), is also seen as an important limitation. Noteworthy is also that financial aspects are not mentioned often. This contrasts with other studies, in which technical issues and financial issues are mentioned in equal measure [12]. The obvious question is then whether money plays an important role according to medical specialists. Finally, these systems should be able to offer support in decision making for diagnosis and treatment [19].

The participants indicated that it is necessary to fulfil some of the preconditions for success before the EMR can make a positive contribution to the hospital's daily practice. That is essentially the source of the medical specialist's resistance. It takes a lot of effort and time to keep the patient file up to date. The only result is a nonpaper file, which medical specialists appreciate but is ultimately not enough to motivate them. It is tempting to make encoding medical data mandatory. However, without interventions in the organization, this is doomed to fail because many medical specialists are unwilling or unable to comply. In the end, inefficiently organized processes will then be automated at great cost and effort, while remaining inefficient at their core [20].

So, prior to the question of how to improve the available processes comes the basic question, What can be improved? Are the available business processes principally accepted, or can the search be directed toward a change in the existing processes [21]?

As stated, it is essential to understand what medical specialists perceive to be key aspects that either support or hinder the use of EMRs to positively impact diagnosis and treatment, now and in the future. These findings may help decide how medical processes can be improved with modern IT. Important in this approach is that the possibilities of modern IT, especially for advanced decision support, be taken as a starting point [22].

Limitations of the Study

Our study had several limitations. We only interviewed medical specialists from hospitals that lacked nursing documentation in 2012-2014. These hospitals found themselves at the lowest stage of EMR use according to the EMRAM model but also had great potential to improve effectively and were able to learn from other hospitals. However, this qualitative study might still have

been too early in the hospitals' implementation of EMRs to identify aspects that are relevant for mature use of EMRs.

The age of the interviewed medical specialists varied between 42 and 62 years. These are medical specialists with a great deal of experience in the field. A question is whether the perspectives of younger medical specialists correspond with the perspectives of their older colleagues. Follow-up research might answer this question.

Conclusions

The 11 medical specialists shared positive experiences of EMR use when comparing it to the formerly used paper records. The fact that the health professionals involved can access patient data at any time they need it is considered important. However, in practice, potential quality improvement lags as long as decision support cannot be applied due to the lack of a fully coded patient record.

Acknowledgments

The authors wish to thank Leonard Witkamp, professor of telemedicine at UMC Amsterdam, for his valuable early review of the article.

Authors' Contributions

RvP planned and conceptualized the study, developed the interview guide, acquired data, analyzed and interpreted data, and drafted the manuscript. GS assisted in developing the interview guide and was involved in interpretating the data, preparing an early version of the manuscript, and revising the manuscript. AB assessed the adapted use of the taxonomy, as published in [15]. KR supervised the study. All authors have read and approved the final manuscript. The data sets generated or analyzed during this study are available, after anonymization, from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no competing interests. No funding was acquired.

References

1. Lin Y, Lin M, Chen H. Do electronic health records affect quality of care? Evidence from the HITECH Act. Inf Syst Res 2019 Mar;30(1):306-318 [FREE Full text] [doi: 10.1287/isre.2018.0813]
2. Wani D, Malhotra M. Does the meaningful use of electronic health records improve patient outcomes? J Oper Manag 2018 Jun 22;60(1):1-18. [doi: 10.1016/j.jom.2018.06.003]
3. van Poelgeest R, Daniels J, de Leeuw RJ, Schrijvers G. Profile of the Digitization of Patient Medical Records in Dutch Hospitals.: HIMSS; 2015 Oct 15. URL: http://www.jhimdigital.org/jhim/fall_2015?pg=38#pg38 [accessed 2021-10-18]
4. van Poelgeest R, Heida J, Pettit L, de Leeuw RJ, Schrijvers G. The association between eHealth capabilities and the quality and safety of health care in the Netherlands: Comparison of HIMSS Analytics EMRAM data with Elsevier's 'The Best Hospitals' data. J Med Syst 2015 Sep 5;39(9):90-96 [FREE Full text] [doi: 10.1007/s10916-015-0274-7] [Medline: 26242750]
5. van Poelgeest R, Groningen GJ, Daniels J, Roes K, Wiggers T. Level of Digitization in Dutch Hospitals and the Lengths of Stay of Patients with Colorectal Cancer.: Springer; 2017 Apr 08. URL: https://link-springer-com.proxy.library.uu.nl/article/10.1007/s10916-017-0734-3 [accessed 2021-10-18]
6. Healthcare Information and Management Systems Society. Electronic Medical Record Adoption Model (EMRAM). 2013. URL: https://www.himss.org/what-we-do-solutions/digital-health-transformation/maturity-models/electronic-medical-record-adoption-model-emram [accessed 2021-10-19]
7. Li R, Niu Y, Scott SR, Zhou C, Lan L, Liang Z, et al. Using electronic medical record data for research in a Healthcare Information and Management Systems Society (HIMSS) Analytics Electronic Medical Record Adoption Model (EMRAM) stage 7 hospital in Beijing: cross-sectional study. J Med Internet Res 2021 Aug 03;9(8):e24405 [FREE Full text] [doi: 10.2196/24405] [Medline: 34342589]
8. Baines R, Langelaan M, de Bruijne M, Spreeuwenberg P, Wagner C. How effective are patient safety initiatives? A retrospective patient record review study of changes to patient safety over time. BMJ Qual Saf 2015 Sep;24(9):561-571 [FREE Full text] [doi: 10.1136/bmjqs-2014-003702] [Medline: 26150548]
9. Adler-Milstein J, Woody Scott K, Jha A. Leveraging EHRs to improve hospital performance: the role of management. Am J Manag Care 2014 Nov;20(11 Spec No. 17):SP511-SP519 [FREE Full text] [Medline: 25811825]
10. Yen P, McAlearney AS, Sieck C, Hefner J, Huerta T. Health information technology (HIT) adaptation: refocusing on the journey to successful HIT implementation. J Med Internet Res 2017 Sep 07;5(3):e28 [FREE Full text] [doi: 10.2196/medinform.7476] [Medline: 28882812]
11. Lakbala P, Dindarloo K. Physicians' perception and attitude toward electronic medical record. Springerplus 2014;3:63 [FREE Full text] [doi: 10.1186/2193-1801-3-63] [Medline: 24516790]
12. Ankem K, Turpin J, Uppala V. Physician adoption of electronic health records: a visualisation of the role of provider and state characteristics in incentive programme participation. 2016. URL: http://informationr.net/ir/21-2/paper715.html#.YW5OkRpByUk [accessed 2021-10-19]

13. Yin RK. Qualitative Research from Start to Finish. New York; London: GUILFORD PRESS; 2011.
14. Pettit L. Understanding EMRAM and how it can be used by policy-makers, hospital CIOs and their IT teams. World Hosp Health Serv 2013;49(3):7-9. [Medline: 24377140]
15. Boonstra A, Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Serv Res 2010 Aug 06;10:231 [FREE Full text] [doi: 10.1186/1472-6963-10-231] [Medline: 20691097]
16. Frankovich J, Longhurst CA, Sutherland SM. Evidence-based medicine in the EMR era. N Engl J Med 2011 Nov 10;365(19):1758-1759. [doi: 10.1056/nejmp1108726]
17. Gagnon M, Simonyan D, Ghandour EK, Godin G, Labrecque M, Ouimet M, et al. Factors influencing electronic health record adoption by physicians: a multilevel analysis. Int J Inf Manag 2016 Jun;36(3):258-270. [doi: 10.1016/j.ijinfomgt.2015.12.002]
18. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018 May 01;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
19. Sutton R, Pincock D, Baumgart D, Sadowski D, Fedorak R, Kroeker K. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
20. Porter M, Lee T. The Strategy That Will Fix Health Care. 2013. URL: https://hbr.org/2013/10/the-strategy-that-will-fix-health-care [accessed 2021-10-19]
21. Badiru AB. Systems Engineering Models: Theory, Methods,Applications. NY: CRC Press; 2019.
22. Jenders RA. Advances in clinical decision support: highlights of practice and the literature 2015-2016. Yearb Med Inform 2017 Aug 19;26(01):125-132. [doi: 10.1055/s-0037-1606493]

Abbreviations

AE: adverse event
CDSS: clinical decision support system
CLMA: closed loop medication administration
DSCA: Dutch Surgical Colorectal Audit
EMR: electronic medical record
EMRAM: Electronic Medical Record Adoption Model
HIMSS: Healthcare Information and Management Systems Society
GP: general physician
MDL: maag-darm-lever
NIVEL: Netherlands Institute for Health Services Research
LOS: length of stay
PC: personal computer
UMC: University Medical Centre
WMO: Wet Medisch-Wetenschappelijk Onderzoek met Mensen

Edited by A Kushniruk; submitted 02.02.21; peer-reviewed by B Van Dort, I Adeleke; comments to author 14.06.21; revised version received 25.07.21; accepted 10.08.21; published 27.10.21.

Please cite as:
van Poelgeest R, Schrijvers A, Boonstra A, Roes K
Medical Specialists' Perspectives on the Influence of Electronic Medical Record Use on the Quality of Hospital Care: Semistructured Interview Study
JMIR Hum Factors 2021;8(4):e27671
URL: https://humanfactors.jmir.org/2021/4/e27671
doi:10.2196/27671
PMID:34704955

Original Paper

Stakeholder Perspectives on an Inpatient Hypoglycemia Informatics Alert: Mixed Methods Study

Nestoras Mathioudakis[1], MHS, MD; Moeen Aboabdo[1], MBChB, MPH; Mohammed S Abusamaan[1], MPH, MD; Christina Yuan[2], MPH, PhD; LaPricia Lewis Boyer[3], CCRP; Scott J Pilla[3], MHS, MD; Erica Johnson[4], MD; Sanjay Desai[5], MD; Amy Knight[4], MD; Peter Greene[6], MD; Sherita H Golden[1], MHS, MD

[1]Division of Endocrinology, Diabetes & Metabolism, Department of Medicine, Johns Hopkins University, Baltimore, MD, United States

[2]Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University, Baltimore, MD, United States

[3]Division of General Internal Medicine, Department of Medicine, Johns Hopkins University, Baltimore, MD, United States

[4]Department of Medicine, Johns Hopkins Bayview Medical Center, Johns Hopkins University, Baltimore, MD, United States

[5]Department of Medicine, Johns Hopkins University, Baltimore, MD, United States

[6]Department of Cardiac Surgery, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Nestoras Mathioudakis, MHS, MD
Division of Endocrinology, Diabetes & Metabolism, Department of Medicine, Johns Hopkins University
1830 E. Monument Street, Suite 333
Baltimore, MD, 21287
United States
Phone: 1 410 502 8089
Fax: 1 410 367 2042
Email: nmathio1@jhmi.edu

Abstract

Background: Iatrogenic hypoglycemia is a common occurrence among hospitalized patients and is associated with poor clinical outcomes and increased mortality. Clinical decision support systems can be used to reduce the incidence of this potentially avoidable adverse event.

Objective: This study aims to determine the desired features and functionality of a real-time informatics alert to prevent iatrogenic hypoglycemia in a hospital setting.

Methods: Using the Agency for Healthcare Research and Quality Five Rights of Effective Clinical Decision Support Framework, we conducted a mixed methods study using an electronic survey and focus group sessions of hospital-based providers. The goal was to elicit stakeholder input to inform the future development of a real-time informatics alert to target iatrogenic hypoglycemia. In addition to perceptions about the importance of the problem and existing barriers, we sought input regarding the content, format, channel, timing, and recipient for the alert (ie, the *Five Rights*). Thematic analysis of focus group sessions was conducted using deductive and inductive approaches.

Results: A 21-item electronic survey was completed by 102 inpatient-based providers, followed by 2 focus group sessions (6 providers per session). Respondents universally agreed or strongly agreed that inpatient iatrogenic hypoglycemia is an important problem that can be addressed with an informatics alert. Stakeholders expressed a preference for an alert that is nonintrusive, accurate, communicated in near real time to the ordering provider, and provides actionable treatment recommendations. Several electronic medical record tools, including alert indicators in the patient header, glucose management report, and laboratory results section, were deemed acceptable formats for consideration. Concerns regarding alert fatigue were prevalent among both survey respondents and focus group participants.

Conclusions: The design preferences identified in this study will provide the framework needed for an informatics team to develop a prototype alert for pilot testing and evaluation. This alert will help meet the needs of hospital-based clinicians caring for patients with diabetes who are at a high risk of treatment-related hypoglycemia.

(JMIR Hum Factors 2021;8(4):e31214) doi:10.2196/31214

KEYWORDS

informatics alert; clinical decision support; hypoglycemia; hospital; inpatient

Introduction

Background

Hypoglycemia is a common occurrence in hospitals and has been linked to poor clinical outcomes and increased mortality [1,2]. Patients with and without diabetes can experience acute hypoglycemic episodes in the hospital, which may result in outcomes ranging from mild distress and patient dissatisfaction with cardiac ischemia, arrhythmias, loss of consciousness, stroke, seizures, and coma [1,3,4]. Of the 8 million and rising number of patients with diabetes who are admitted to hospitals in the United States annually, up to 25% may experience a hypoglycemic episode during hospitalization [5]. Approximately half of these episodes can be explained by an underlying illness, such as severe sepsis, renal failure, liver failure, or malignancy [3]; however, the remaining half are iatrogenic in nature, usually resulting from insulin treatment [6].

Insulin remains the recommended therapy for most patients with diabetes during hospitalization [7,8]. Recent data from the Medicare Patient Safety Monitoring System found that glucose-lowering medications were associated with the highest rates of adverse outcomes for all drugs used in the hospital [9]. As insulin accounts for the vast majority of hypoglycemic events [10,11], the Joint Commission and the Institute for Safe Medication Practices have designated it a *high-alert* medication [12]. Insulin is typically administered as a continuous infusion for critically ill patients in the intensive care unit (ICU) and as subcutaneous injections in noncritically ill patients in general medical or surgical wards. In contrast to the ICU setting where insulin adjustments are driven by nurse-managed protocols, insulin titration in the non-ICU setting is prescriber-driven and requires evaluation of a complex set of clinical, laboratory, and pharmacological parameters.

Clinical decision support (CDS) tools in electronic medical record (EMR) systems have been increasingly used in the United States and have been shown to improve the processes of care and clinical outcomes [13-15]. These tools have been used to alert clinicians, suggest diagnostic or treatment recommendations, and provide contextually pertinent information to optimize care for a wide variety of indications in the hospital setting, ranging from sepsis to acute kidney injury [16-18]. It stands to reason, then, that CDS could be used to improve care of patients vulnerable to hypoglycemia. Prediction models using large EMR data sets have been developed to trigger alerts in patients at risk of hypoglycemia [19-22]. Several studies have found that these predictive models can decrease the rate of inpatient hypoglycemic episodes, with one review finding that it could decrease the incidence of severe hypoglycemia by up to 68% [19,23].

At our institution, we have several existing CDS tools to guide clinicians in selecting a safe and effective initial insulin dosing regimen, including a mandatory subcutaneous insulin order set and an optional insulin CDS tool [24]. Although our subcutaneous insulin CDS tool was derived from evidence-based basal-bolus insulin dosing protocols [25], even universal use of this tool would not be expected to prevent all iatrogenic hypoglycemic events in the hospital because of the impact of acute illness and other clinical factors on glucose homeostasis (eg, nutritional status, renal function, and steroid doses).

Objectives

To address the need for real-time hypoglycemia risk detection, we recently developed a machine learning algorithm using EMR data that accurately predicts iatrogenic hypoglycemia in rolling 24-hour windows following each blood glucose reading during hospitalization [22]. In planning the translation of this machine algorithm into a real-time informatics alert, we sought to obtain feedback from inpatient clinicians who are responsible for the day-to-day management of blood glucose in the hospital. The objective of this study is to obtain stakeholder input regarding the design of a real-time hypoglycemia informatics alert for use in a hospital setting.

Methods

Study Design

This was a mixed methods study that integrated quantitative and qualitative data. The initial stage consisted of an electronic survey (administered in February 2019) that was sent to physicians and advanced practice providers at 2 academic medical centers located in Baltimore, Maryland. The second stage consisted of 2 separate focus groups (conducted in September 2019) with participants recruited from the pool of survey respondents, the goal of which was to expand further on the answers to the survey and identify common themes. This study was approved by the institutional review board at the Johns Hopkins School of Medicine. For the electronic survey, consent was implied from respondent completion, and written informed consent was obtained from all participants in the focus group sessions.

Survey

To evaluate the key features and functionality of a hypoglycemia risk alert, we developed a 21-item electronic survey that was administered through SurveyMonkey (SVMK Inc) via an embedded email hyperlink (Multimedia Appendix 1). The survey required approximately 7-10 minutes to complete and was sent to hospital-based clinicians involved in glucose management, including medical and surgical residents, hospitalists, surgical advanced practice providers, and inpatient diabetes nurse practitioners (NPs). To encourage participation, the email invitation with the survey link was sent directly by residency program directors (medicine, surgery, neurology, obstetrics/gynecology [OB/GYN]) to all residents in their programs, and by program leaders in the hospitalist, surgical advance practice provider groups, and inpatient diabetes management service. Thus, the total number of recipients who received the survey link (and hence the response rate) was not known by the study investigators.

Survey questions were developed using the Agency for Healthcare Research and Quality Five Rights of effective CDS as a guiding framework: the right *information* to the right *person*, in the right interventional *format*, through the right *channel*, and at the right *time* in the workflow [26]. *Right information* refers to the content and presentation of information to the end user (accuracy, estimated risk and reasons for predicted risk, and recommended action). *Right person* refers to the member of the health care team who is most appropriate to receive and respond to the alert and the method of identifying this individual. *Right format* refers to the type of CDS used to address the clinical scenario and *right channel* refers to the platform for communicating the alert. We considered several formats or channels within and outside our EMR system (Epic Version 2018, Epic Systems Corporation), including the best practice advisory (BPA), glucose management print group report, patient system lists, patient header, InBasket messaging, and text messaging. A description of each of these formats or channels is provided in Table 1, with example screenshots in Figure 1. Finally, *right time in the workflow*, a critical component of successful CDS interventions, refers to the time when the alert is presented to the end user in their usual clinical workflow to minimize disruption and achieve desired action. Multimedia Appendix 2 categorizes the survey questions according to the *Five Rights* topics.

Table 1. Description of formats and channels considered for alert.

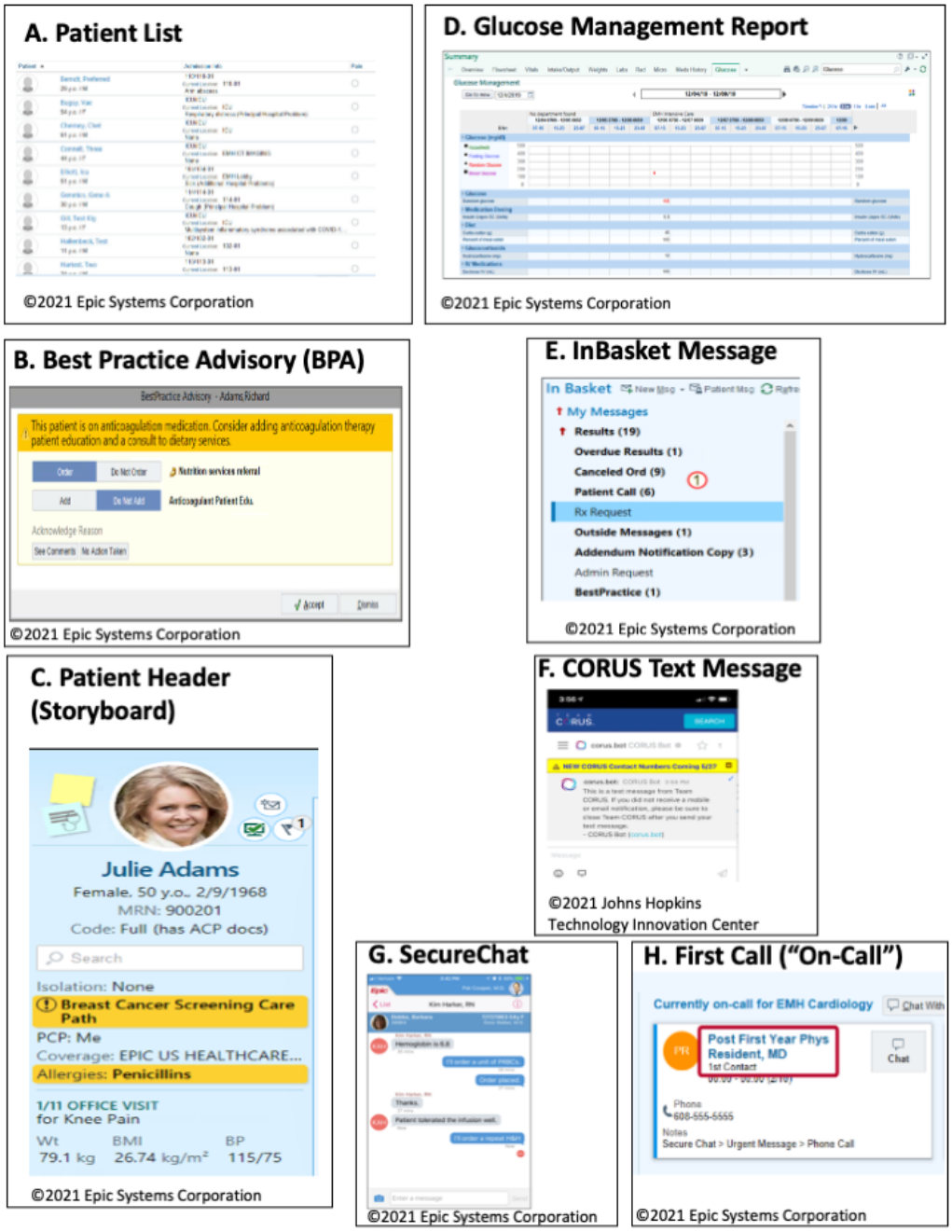
Term	Description	Epic best practice recommendation
Patient list (Figure 1A)	Central hub for clinicians to see patients in their unit and across facility. System list can be compiled across hospital or providers can add a column in their existing patients' lists to identify patients who meet certain criteria.	Identify populations of patients that clinicians need to review regularly or notify clinicians of individual patients who need their attention.
BPA[a] (Figure 1B)	Alert that appears based on a wide variety of events and actions.	Use for one-time events that do not recur on a regular basis. Restrict how often and to whom BPAs appear so that they appear only to users who can act on them at a time when they perform the action. Limit the number of BPAs that appear in a separate window.
Patient header (Storyboard; Figure 1C)	Provides patient information relevant to user's specific role along left side of screen. BPAs can appear in Patient Header so they do not interrupt a clinician's workflow.	Show information to clinicians that they need to review or that should be available at a glance from anywhere in the chart.
Glucose management report (Figure 1D)	Glucose management report contains summary of all subcutaneous insulin doses over previous 24 hours, and all glucose and insulin doses administered since admission. Allows user to review relevant information about the patient from one spot in EMR[b] as opposed to searching several areas to compile information.	Use to show information that a clinician needs to make decisions based on the total information compiled in the report.
InBasket message (Figure 1E)	Secure, closed, task-based messaging system to send and receive information about patient care, directly linking messages to patient's accounts, chart, laboratory results, and orders.	Good fit for questions or issues that do not need to be handled immediately because users might not regularly check all of their InBasket messages.
CORUS text message (Figure 1F)	CORUS: Secure text messaging system developed by Johns Hopkins Technology Innovation Center allowing users to communicate within channels or groups on computer or mobile device. This communication channel is external to EMR.	N/A[c]
Secure chat text message (Figure 1G)	Epic Secure Chat (deployed after completion of this study and will ultimately replace existing CORUS text messaging system) allows users to have conversations with a single recipient or with a group of colleagues securely on a mobile device. This communication channel is internal to EMR.	Intended for quick coordination between members of the team
First Call (Figure 1H)	Designated field to specify on-call provider in EMR.	N/A

[a]BPA: best practice advisory.

[b]EMR: electronic medical record.

[c]N/A: not applicable.

Figure 1. Screenshots examples of proposed formats and channels for informatics alert.



Focus Groups

To gain a more in-depth understanding of the perceptions, preferences, and perceived barriers to a hypoglycemia informatics alert, we conducted 2 focus group sessions, each consisting of 6 different participants, after the results of the survey were analyzed. Hospital-based physicians, NPs, and PAs were eligible for participation. Participants who responded to the initial survey expressing interest in further research related to the topic were invited to the focus group sessions. No participant identifiers were used during the session, and all participants were referred to by a given letter at the start of the session.

The focus group sessions lasted 60-90 minutes and were led by an experienced focus group moderator (LLB). An institutional review board approved–approved structured interview guide was used (Multimedia Appendix 3), with questions designed to delve into survey results and reasoning behind responses. The principal investigator (NM) was also presented to address or clarify any participant questions during the session. The sessions were audio-recorded and professionally transcribed. The investigators further reviewed and corrected the transcriptional errors. The participants were provided with food during the session and a US $100 gift card.

Data Analysis

Descriptive statistics were used to summarize the characteristics of the survey respondents and focus group participants. For binary, multiple-choice responses and Likert response data, the number and percentage of responses were provided. For rank items, the average ranking was calculated as the weight of the ranked position. MAXQDA (Verbi Software, 2019) was used

for qualitative analysis and thematic coding of data from the focus group sessions. We used a combination of deductive and inductive methods in our analyses [27]. Before coding the data, we identified an initial set of deductive codes centered on the CDS Five Rights framework. Close reading of the transcripts then informed the development of inductive codes, reflecting *the ground*, or participants' experiences. When the code structure was considered final (ie, no new concepts were apparent), one researcher independently applied the finalized code structure and synthesized the data into patterns (ie, a cohesive category of responses found across our participants) and themes (ie, a broad concept or topic that aggregates patterns).

Results

Study Participants

Table 2 shows the characteristics of the survey respondents and participants in the 2 focus group sessions. There were a total of 102 survey respondents, most of whom were internal medicine (IM) physicians, with fairly even representation from trainees and faculty or staff. Focus group 1 comprised 6 NPs (of whom 4 were inpatient diabetes specialists) and Focus group 2 comprised 6 physicians (half of whom were medicine residents).

Table 2. Characteristics of survey and focus group respondents.

Characteristic	Survey	Focus group 1	Focus group 2
Participants, n	102	6	6
Provider type, n (%)			
Physician	73 (71.6)	0 (0)	6 (100)
Nurse practitioner	21 (20.6)	6 (100)	0 (0)
Physician assistant	6 (5.9)	0 (0)	0 (0)
Other	2 (1.9)	0 (0)	0 (0)
Specialty, n (%)			
Medicine	52 (50.9)	0 (0)	3 (50)
Surgery	26 (25.5)	2 (33.3)	0 (0)
Endocrinology or diabetes	15 (14.7)	4 (66.7)	1 (16.6)
Neurology or neurosurgery	5 (4.9)	0 (0)	1 (16.6)
Obstetrics/gynecology	3 (2.9)	0 (0)	1 (16.6)
Other	1 (0.9)	0 (0)	0 (0)
Level of training, n (%)			
Resident or fellow	55 (53.9)	0 (0)	6 (100)
Faculty or staff	47 (46.1)	6 (100)	0 (0)
Age (years), mean (SD)	—[a]	39.8 (3.9)	29.5 (2.1)
Experience (years), n (%)			
<5	—	1 (16.6)	6 (100)
5-10	—	1 (16.6)	0 (0)
≥10	—	4 (66.7)	0 (0)

[a]Data not collected in the survey.

Survey Responses

Figure 2 summarizes the responses to the survey questions related to the importance of the problem and the perceived benefit of the proposed alert. There was unanimous agreement that preventing insulin-related hypoglycemia in hospitalized patients is an important priority, and 47% (48/102) either agreed or strongly agreed that preventing insulin-related hypoglycemia in the hospital is challenging. Nearly all providers reported that they reviewed blood glucose data on insulin-treated patients and almost all providers felt that they recognized glycemic patterns indicative of a need to adjust insulin doses. Accordingly, 82.3% (84/102) of the respondents felt confident in their ability to safely adjust their insulin doses. Over two-thirds felt that a real-time hypoglycemic alert would be beneficial.

Figure 2. Survey results related to importance of problem and perceived benefit of informatics alert.

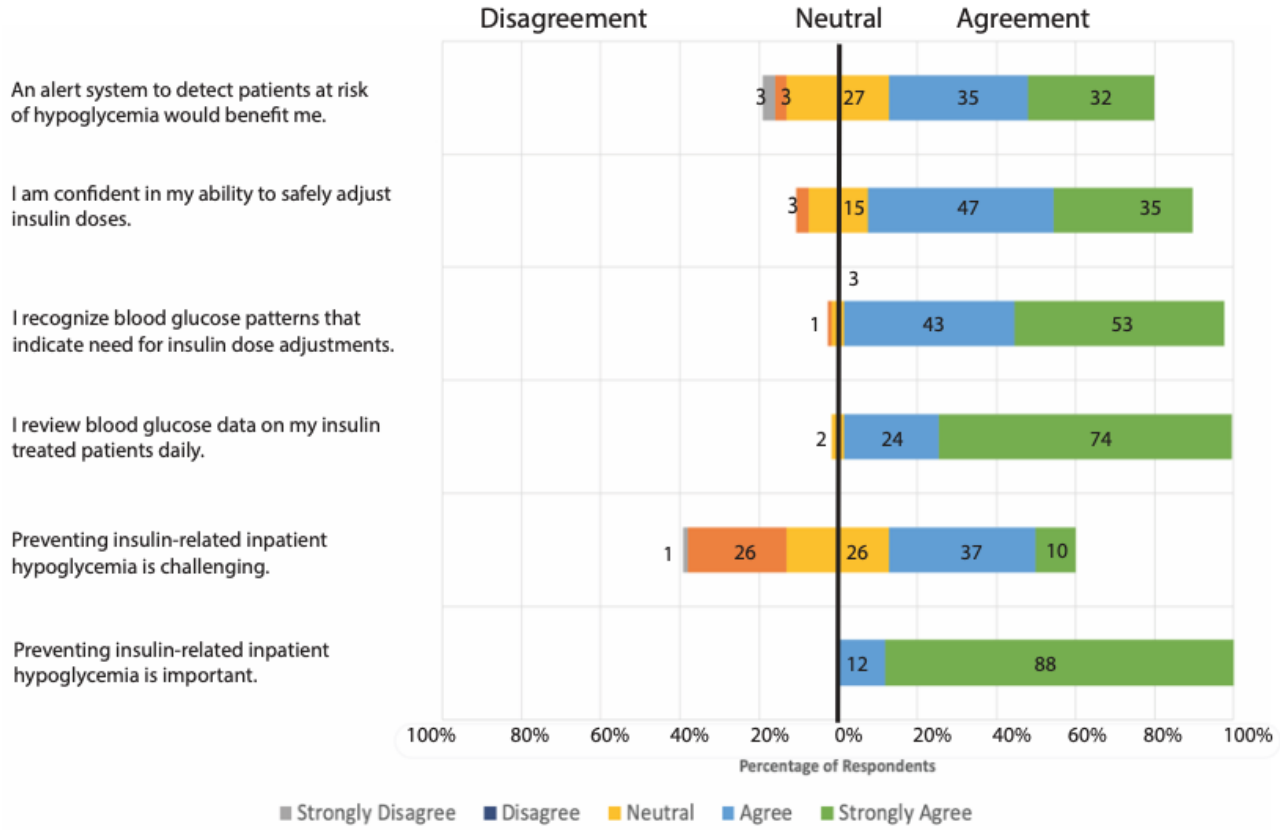


Table 3 summarizes the survey responses related to the proposed information alert. The top 3 preferred formats were an alert indicator in the patient header, secure text message, and alert prompt displayed in the existing glucose management report. The least desired formats were an Epic InBasket message, a patient system list, and a pop-up BPA. In terms of information included in the alert, 78.4% (80/102) of the respondents indicated that they would want to know the specific reasons why an individual patient is predicted to be at risk based on the prediction model; 64.7% (66/102) of the respondents wanted the alert to categorize the predicted hypoglycemia risk into low, medium, and high levels; and 62.8% (64/102) wanted the alert to give an estimated probability of hypoglycemia. A minority of respondents (21/104, 23.5%) indicated that they would want the alert to hyperlink to an actual prediction model, and 16.7% (17/102) indicated that they would want the validated accuracy of the model displayed in the alert. With respect to accuracy, there was a wide range of acceptable sensitivity and specificity thresholds among respondents, with most considering 70% and above to be acceptable for both.

Table 3. Survey results.

Survey question	Values
What is your preferred format for the tool? (average rank, higher score=more desirable)	
Patient header	4.59
Text message	4.03
Glucose management report	3.81
BPA[a]	3.65
Patient list	2.94
Epic InBasket message	1.98
What piece of information would you like to see included in the real-time alert? (select all that apply; n=102), n (%)	
Specific reason or reasons a patient is at risk based on the prediction model	80 (78.4)
Categorized risk of hypoglycemia (eg, low, medium, high)	66 (64.7)
Patients' estimated probability of hypoglycemia	64 (62.8)
Hyperlink to actual prediction model	24 (23.5)
Validated accuracy of the model	17 (16.7)
None of the above is necessary	3 (2.9)
What feature would you like to see incorporated in the real-time alert? (select all that apply; n=102), n (%)	
Recommended action	90 (88.2)
Ability to acknowledge the alert	61 (59.8)
Ability to ignore or override alert	61 (59.8)
Direct link to subcutaneous insulin order set	54 (52.9)
Direct link to subcutaneous insulin decision support tool	29 (28.4)
Ability to consult endocrinology or inpatient diabetes management service	43 (42.2)
None of the above	1 (0.9)
Should the real-time informatics alert trigger an endocrinology consult? inpatient diabetes management service consult? (n=101), n (%)	
No	85 (84.2)
Yes	16 (15.8)
What is your preferred channel for the alert tool? (select all that apply; n=102), n (%)	
CORUS text message	63 (61.8)
BPA tool	52 (50.9)
Other channels	15 (14.7)
Epic inBasket message	5 (4.9)
When would you like to receive the alert in your workflow? (single choice; n=102), n (%)	
As soon as hypoglycemia risk is detected	47 (46.1)
At the same time everyday	27 (26.5)
When opening the EMR[b] of a patient predicted to be high risk	24 (23.5)
Other	4 (3.9)
Who should receive the alert? (select all that apply; n=102), n (%)	
Person listed as *first call* on the EMR	86 (84.3)
Nurse	47 (46.1)
Attending physician	15 (14.7)
Clinical nurse specialist (nurse practitioner)	15 (14.7)
What is lowest sensitivity you'd consider clinically acceptable for the proposed alert tool? (single choice; n=101), n (%)	

Survey question	Values
50%-59%	3 (2.9)
60%-69%	10 (9.9)
70%-79%	37 (36.6)
80%-89%	37 (36.6)
90%-100%	14 (13.9)
What is lowest specificity you'd consider clinically acceptable for the proposed alert tool? (single choice; n=100), n (%)	
50%-59%	3 (3)
60%-69%	6 (6)
70%-79%	30 (30)
80%-89%	41 (41)
90%-100%	20 (2)
Are you interested in participating in a clinical design team to build the informatics alert? (n=96), n (%)	
No	73 (76)
Yes	23 (24)

[a]BPA: best practice advisory.

[b]EMR: electronic medical record.

The vast majority of the respondents (90/102, 88.2%) indicated that they would want the alert to provide some recommended action, and 59.8% (61/102) of respondents suggested that there should be an option acknowledge the alert or ignore or override the alert. Other features of the alert desired by approximately half or less of the respondents included direct hyperlinks to our subcutaneous insulin order set (54/102, 52.9%) and subcutaneous insulin decision support tool (29/102, 28.4%) and the ability to directly consult endocrinology or inpatient diabetes management service (43/102, 42.2%). Although some respondents felt that the ability to easily consult endocrinology or diabetes services was important, most (85/101, 84.2%) felt that the alert should not automatically trigger a consult.

Regarding the *right person* to receive the alert, 84.3% (86/102) of respondents indicated the person listed as *first call* in the EMR should be notified. In addition to the *first call* provider, 46.1% (47/102) of respondents indicated that the patient's nurse should also be notified. With respect to the workflow, 46.1% (47/102) of respondents indicated they would want to be notified as soon as the hypoglycemia risk was identified, whereas 26.5% (27/102) preferred the same time every day and 23.5% (24/102) preferred being notified only when entering the chart of a relevant patient.

Focus Groups

The core themes discussed during the focus groups centered on the CDS five rights format, with the aim of receiving feedback on how a functional alert system could best suit front-link clinicians at a major hospital, as well as anticipated barriers and challenges. Representative quotes illustrating each key theme are presented in Table 4.

Table 4. Representative quotes from focus groups.

Theme and pattern	Representative quotes
Right information	
Accuracy	• "It's very important when it's first deployed that the alert is highly accurate...Because otherwise, I think you risk people developing an attitude that they're not going to pay attention to it." [OB/GYN[a] resident] • "I think less accurate alert risks developing provider fatigue." [OB/GYN resident] • "As we get more data over time, you can include something about, like, the probability of a hypoglycemic event in your patient is 53% of all patients with this probability in the past X amount of time. What number or percentage went on to have a hypoglycemic event?" [OB/GYN resident]
Trigger	• "I would want to know why the alert was triggered, initially. It's not that as a provider that I don't trust what the computer has calculated to be whatever algorithm that is coming out for my patient, but for my own education and learning." [OB/GYN resident] • "I think having an option that says this alert is inaccurate and not only having that option but that triggering someone to review that alert and see why it was triggered and hopefully, revise it so that it's more accurate." [IM[b] resident]
Recommended action	• "I would say that it would be helpful if the options were there for you to check...if like, you know-like D5 and D10 infusions were there as an option to check to make them an active order." [Surgical NP[c]] • "Maybe putting something like...reduce [insulin dose] by 20% or something that would be helpful." [Diabetes NP] • "One of the things I really appreciate about TREWS [sepsis alert] is that it will guide you through the algorithm and the criteria you need to meet in order to treat the sepsis." [IM resident]
Right person	
Nurses	• "I think the nurse should be one of the first (to be contacted)." [Diabetes NP] • "I think the way the TREWS [sepsis alert] is set up is that the nurse gets the first alert and they're responsible for contacting first call or whoever the primary team is. And so, the benefit of that is they will know how that service is set up whether or not they use the first call." [IM chief resident]
First call provider	• "First call person...if that's the updated one." [Diabetes NP] • "First call system and then the BPAs[d], as long as we can, kind of, minimize it so it's not—again, the issue with alert fatigue." [IM resident]
Consultant	• "I'm forever getting alerts from things that are not a consultant team's responsibility." [Neurology resident]
Right format	
Laboratory results	• "If you can write something under labs. Because I feel like labs, everyone watches." [Diabetes NP] • "Adding a symbol or something to an actual glucose result to say this person is at risk." [OB/GYN resident]
Glucose management report	• "I think I find that tool to be the one that I use the most when I'm managing a patient's glucose and insulin because it gives me a good way to review a 24-hour snapshot." [OB/GYN resident] • "I really like the glucose management report. I think for internal medicine, we use it...all the time." [IM resident] • "I think it would be great to have it there. I don't think it's mutually exclusive from the other ones, but I will say for an alert, people are not always going to look in there at the right time that you want them to know about the information. This raises the concern that although the tool is made to help specifically with glucose management, that a lack of use by some staff may lead to delay in response to an alert." [IM chief resident]
BPA	• "BPAs...interrupt your workflow so much." [IM chief resident]
Epic InBasket	• "Epic InBasket, I agree...It's useless. Nobody's going to look there for urgent things." [IM chief resident] • "Especially the EPIC InBasket, we get so many messages every day. Results to follow-up on. It would just get lost. I mean, there's no way I would ever see that." [Endocrinology fellow]
Patient header	• "Perhaps there's something in the header that tells you need to go look at that tab." [OB/GYN resident] • "I think, actually, the header might be because—like with TREWS [sepsis alert], it's not interfering. It doesn't, like, get in your face and make you answer something, but it's there." [IM resident] • "The BPA thing that will show up in the header. It'll just list that you have BPAs that need to be addressed." [IM chief resident] • "I feel like our patient headers are very crowded, currently, with quite a bit of information. And so, I'm not entirely sure the best format or buildout to make something appear in the header. Maybe it could be something like a symbol that appears that then indicates that you should go look at the glucose management tab." [OB/GYN resident]

Theme and pattern	Representative quotes
Right channel	
CORUS (text messages)	• "All of us, I think, as housestaff, also check CORUS pretty religiously." [IM resident] • "I think in terms of being alerted to it physically as a house staff member, CORUS is the best." [IM resident]
Right time in workflow	
Real time	• "Alerts should be time sensitive and in real time." [Surgical NP]
"Snooze" feature	• "What do you think about the idea of—you know, when your computer lets you know there are updates and it needs to be restarted and you say, 'Not now. Try again in an hour. Not now. Try again tonight.' I'm saying this alert comes up—I'm opening a chart for a specific reason. I have a task or multiple tasks in mind." [OB/GYN resident] • "Or if there was a way for it to be paused and then it pops up again when you go to click out of the chart. It's like, hey, don't forget—you've got this thing to do." [Neurology resident]
Disruption	• "Oftentimes, you'll just click something to get it out of the way, do what you're doing, and then you'll forget about it afterward." [IM chief resident] • "Which I think would be helpful to, like-I need to put in another order right now that's actually more urgent for the patient, believe it or not. And I don't want to forget to do that, because I'm messing around now in their Lantus [insulin] dosing, and there and then I forget...or I now have gotten into this big rabbit hole of 6 different orders that I have to place in order to put in the original order that I wanted to put in." [Surgical NP]
Barriers or challenges	
Importance of problem	• "I think the good question to ask is whether or not we have enough of a problem in which patients get into real trouble, as opposed to just having a glass of orange juice." [Surgical NP] • "I think providers are more reactive to hyperglycemia than they are hypoglycemia." [Diabetes NP] • "And that's rare, right? 99% of these patients are treated with a glass of juice." [Surgical NP]
Communication	• "Plus, there are so many other providers involved when you order something, or you recommended something." [Diabetes NP]
Provider factors	• "When they see the blood sugar is high, [providers]...keep giving insulin without understanding the duration of action of the insulin. So, the patient ends up getting stacked." [Diabetes NP]
Patient factors	• "It's impossible for us to walk in and put the tray down in front of them, check their sugar, give them insulin, and then they are guaranteed to eat >50% of the tray in front of them." [Surgical NP] • "How much different a diet, in particular, relates to this. People would be on like 100 units of insulin a day at home and if you put them on that [amount of insulin] here, [their blood glucose] will shoot to 0." [IM resident] • "Our patients are non-compliant people, and so, they come in and their home regimen has been ramped up in the outpatient setting because they just aren't doing it and then you're trying to guess, sort of, like what are your actual insulin needs and you either become too conscientious and they're entirely way too hyperglycemic or we're not conscientious enough in trying to guess, sort of, what's their appropriate doses. It's really challenging." [Neurology resident]
Alert fatigue	• "As a consulting physician, to be honest with you, I don't really even look at them. I just kind of click to get it out of my way button...I'm not the one who actually has to deal with it 90% of the time." [Endocrinology fellow] • "What drives me most crazy is, if you've already answered the questions and then the next time you log into Epic, it shows up again, and again, and again, and again." [Surgical NP] • "Then I know we have the new hypoglycemia alerts that pop up, but I think they pop up very, very, very frequently to the point that I think it's almost starting to cause a little bit of fatigue." [IM resident]

[a]OB/GYN: obstetrics/gynecology.

[b]IM: internal medicine.

[c]NP: nurse practitioner.

[d]BPA: best practice advisory.

Right Information

Accuracy was a common theme. According to the surveys, most of our participants wanted the alert to be at the very least 70% sensitive and specific. One participant noted as follows:

> *It's very important when it's first deployed that the alert is highly accurate...Because otherwise, I think you risk people developing an attitude that they're not going to pay attention to it.* [OB/GYN resident]

Participants raised concerns that if the specificity was low at launch, the alert would go off repeatedly and would be ignored.

Many noted that similar alert systems have been deployed at our institution before validation of accuracy; as a result, many of the alerts were initially unsuccessful at gaining end user buy-in and changing behavior. To avoid this, one participant noted that the algorithm should be continuously adjusted to increase accuracy over time.

A common theme among focus group participants was the importance of defining specific reasons that triggered an alert. Participants felt that sharing this information would present a learning opportunity for end users and could indirectly modify behavior in their care of future patients. Participants also voiced the concern that there should be an option to report the alert if it is inaccurate:

> *I think having an option that says this alert is inaccurate and not only having that option but that triggering someone to review that alert and see why it was triggered and hopefully, revise it so that it's more accurate.* [IM resident]

Finally, echoing the results of our survey, the focus group participants emphasized their desire to have an alert to recommend the appropriate action. The tendency of alert systems to interrupt workflow without recommending an appropriate action leads to clinicians feeling overwhelmed. Participants noted that the best part of the other alert systems was the ability to help with workup and management. One participant noted as follows:

> *One of the things I really appreciate about TREWS [sepsis alert] is that it will guide you through the algorithm and the criteria you need to meet in order to treat the sepsis.* [IM resident]

Right Person

There were mixed opinions among the focus group participants as to the right person to be notified of the alert. Many believed that contacting the individual or service listed as *first call* in the EMR would be the most sensible approach. However, many voiced the concern that the only way a first-call system would be efficient was if the EMR was accurately updated, which may not always be the case. Other participants voiced that the right person to be alerted should always be a nurse. One participant noted that with other alert systems, the nurse is responsible for triaging the alert and determining whether contacting the provider is required:

> *I think the way the TREWS [sepsis alert] is set up is that the nurse gets the first alert and they're responsible for contacting first call or whoever the primary team is. And so, the benefit of that is they will know how that service is set up whether or not they use the first call.* [IM resident]

Right Format

There was substantial heterogeneity in preferences regarding the alert format, and our discussion focused on the pros and cons of each format. Focus group participants noted that an alert indicator in the patient header would be a very suitable place for a hypoglycemia alert; however, concerns were raised that the patient header is already a very crowded space:

> *I feel like our patient headers are very crowded, currently, with quite a bit of information. And so, I'm not entirely sure the best format or buildout to make something appear in the header.* [OB/GYN resident]

Others really valued the idea of using the glucose management report, which is a tabular report summarizing insulin and glucose data in a temporal way that facilitates pattern recognition and insulin dose adjustments but cautioned that this passive approach may not achieve the desired action. A suggestion raised during the focus groups session that was not considered in the electronic survey was made to place an alert symbol next to a patient's glucose laboratory result to notify them of impending risk; this alert would be distinct from an abnormal laboratory result value to distinguish predicted risk from overt hypoglycemia. Participants felt that a *results flag* would be the most effective in increasing situational awareness:

> *If you can write something under labs...because I feel like labs, everyone watches.* [Diabetes NP]

A consensus was also reached that BPAs were not ideal for an alert system, as they were noted to be very interruptive by some of the participants:

> *BPAs...interrupt your workflow so much.* [IM resident]

The Epic InBasket format was dismissed by most participants, as this is not a format routinely used for communication in the inpatient setting:

> *Epic InBasket, I agree...It's useless. Nobody's going to look there for urgent things.* [IM chief resident]

Right Channel

The CORUS text messaging system was the unanimously desired channel for receiving the alert. All participants agreed that it was the quickest and most efficient way to be notified. A few even mentioned that they were too heavily reliant on CORUS, but noted it was the appropriate app to get the staff's attention:

> *All of us...as housestaff also check CORUS pretty religiously.* [IM resident]

Right Timing

The majority agreed that the alert should be direct and in real time to allow for appropriate action. One concern that many had was the fact that the alert might go off in a time when the staff member could not respond. One participant offered the suggestion of a *snooze button* to address this concern:

> *What do you think about the idea of—you know, when your computer lets you know there are updates and it needs to be restarted and you say, "Not now. Try again in an hour. Not now. Try again tonight." I'm saying this alert comes up...I'm opening a chart for a specific reason. I have a task or multiple tasks in mind.* [OB/GYN resident]

Barriers to Alert Systems

The focus group sessions elicited feedback about barriers and challenges related to hypoglycemic prevention and the proposed informatics alert. Some of the themes that emerged were: (1)

iatrogenic hypoglycemia is not serious enough to justify an alert, (2) patient factors that cannot be controlled (eg, snacking without knowledge of the treating team or resistance to taking recommended insulin doses), (3) provider factors, including knowledge gaps, and, most importantly, (4) concerns regarding alert fatigue. Although the survey respondents almost unanimously agreed that preventing hypoglycemia was an important problem, some of the focus group participants seemed to question its significance. For example, one participant commented as follows:

> *I think the good question to ask is whether or not we have enough of a problem in which patients get into real trouble, as opposed to just [needing to treat with] a glass of orange juice.* [Surgical NP]

Many participants voiced concern that the alert system must be structured to prevent alert fatigue. Participants noted that other alert systems were constantly popping up inappropriately leading to *people clicking through them* and interrupting the workflow.

Discussion

Principal Findings

Overall, the findings from this study indicate that iatrogenic hypoglycemia was perceived as an important problem that could potentially be addressed with a real-time informatics alert. We identified several key design requirements and preferences for the alert: (1) sufficiently accurate, (2) recommendations for care, (3) nondisruptive, and (4) communicated in near real-time to the responsible clinician. Several EMR tools, including an alert indicator embedded in the patient header, glucose management report, and flagged laboratory results, were deemed acceptable formats for consideration of hypoglycemia alerts. Although clinicians indicated that pop-up (interruptive) BPAs would be considered unacceptable because of alert fatigue and disruption of clinical workflow, we were surprised by the acceptability of the use of the hospital's text messaging system as a channel of communication for the alert because real-time text messaging could be interrupted.

Comparison With Prior Work

To our knowledge, there are no previous studies soliciting stakeholder input regarding inpatient hypoglycemia informatics alerts. A systematic review and meta-analysis of computerized CDS systems (CDSS) found that push notifications, the ability to execute action, and an interruptive element are the most prevalent features of CDSS that have been evaluated in controlled clinical trials [14]. Absolute incremental improvements in clinical processes or outcomes were demonstrated for CDSS that offered the ability to execute the desired action, and a trend toward improvement was shown for those that allowed for acknowledgment, were behavior-targeted, embedded in the EMR, and interrupted [14].

We recently developed a machine learning algorithm using EMR data that would be considered sufficiently accurate (sensitivity and specificity of approximately 80% and c-statistic of 0.90) based on the input received from stakeholders in this study [22]. As we progress toward translating our machine learning model into a real-time alert, we will seek to consider the key findings from this study to optimize end user buy-in and adoption. Specifically, we intend to develop a near real-time alert that provides specific and actionable recommendations to the ordering provider without disrupting their clinical workflow. The top candidate CDS formats we will consider based on the results of this study are a nonintrusive BPA embedded in the patient header, laboratory result, or existing glucose management report. The selection of the final CDS format will depend in part on our institutional requirements and programing feasibility.

Limitations

It is important to note that there have been three key system changes at our institution since we completed this study, which may have affected our findings. First, our text messaging system (CORUS) has been replaced by an alternative text messaging platform that resides within the EMR and is accessible through mobile devices (SecureChat). The main difference in the text messaging platform change is that the communication resides within the EMR and is accessible via a mobile app. Communications delivered via SecureChat would allow the user to quickly access other relevant information within the patient's chart without the need to access an external system. Second, a new feature called *Storyboard* was released in an Epic software upgrade, which shifted the patient header information from the top of the chart to the left side of the screen. The Storyboard includes patient identifiers and key information, including allergies, infection or isolation, and best practice advisories. At our institution, the Storyboard is already being used to summarize active alerts for a given patient (eg, sepsis) and would be a potential consideration for the development of hypoglycemia alerts. We believe that the inferences of this study related to the patient header would apply to the new Storyboard format. A third system change at our institution has been made in which all services are required to either list an individual or a service pager name (for surgical services) in the first call field. It is expected that the first call field will serve as the source of truth in identifying the responsible provider for a given patient at any given time.

Conclusions

Alert fatigue is a commonly cited limitation in CDSS [28-30]. Despite the fact that avoidance of alert fatigue emerged as one of the most important themes in our analysis, some incongruencies were identified among participants in this study. On the one hand, clinicians strongly opposed a pop-up style BPA; on the other hand, real-time alerting via text messaging was identified as a preferred channel of communication regarding hypoglycemia risk. Although a text message can perhaps be more easily ignored than a pop-up BPA, both forms of alerting could distract the clinician from their current clinical work. On the basis of the overall input we received from our stakeholders, we place greater value on avoidance of alert fatigue and disruption in clinical workflow than immediate notification of the risk of hypoglycemia. Therefore, we will explore several strategies to minimize alert fatigue, such as communicating increased risk in a nonintrusive fashion in the patient header, laboratory result section, or glucose management report.

Future research will assemble an informatics design team to develop prototypes using the most desired formats proposed by stakeholders in this study, and to test these prototypes in a pilot observational study before being widely implemented and evaluated for effectiveness. Ensuring that the alert system follows the CDS Five Rights and adheres and aligns with stakeholder preferences from this study will hopefully improve the usability, adherence, and efficacy of the hypoglycemic alert.

Acknowledgments

NM and MSA made significant contributions to the conception and design of the study, conducting interview sessions, statistical analysis, and interpretation of the results. NM, MA, MSA, and CY drafted the manuscript. LLB led to interview sessions. SG contributed to the study concept, design, and critical revision of the manuscript. All authors reviewed and edited the manuscript. NM and MSA were supported by a grant #K23DK111986 from the National Institute for Diabetes and Digestive and Kidney Diseases.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey question response types and concepts.
[DOCX File , 16 KB - humanfactors_v8i4e31214_app1.docx]

Multimedia Appendix 2
Electronic survey.
[DOCX File , 854 KB - humanfactors_v8i4e31214_app2.docx]

Multimedia Appendix 3
Structured interview guide.
[DOCX File , 23 KB - humanfactors_v8i4e31214_app3.docx]

References

1. International Hypoglycaemia Study Group. Hypoglycaemia, cardiovascular disease, and mortality in diabetes: epidemiology, pathogenesis, and management. Lancet Diabetes Endocrinol 2019 May;7(5):385-396. [doi: 10.1016/S2213-8587(18)30315-2] [Medline: 30926258]
2. Tunbridge W. Factors contributing to deaths of diabetics under fifty years of age. On behalf of the Medical Services Study Group and British Diabetic Association. Lancet 1981 Sep 12;2(8246):569-572. [doi: 10.1016/s0140-6736(81)90950-8] [Medline: 6116012]
3. Brutsaert E, Carey M, Zonszein J. The clinical impact of inpatient hypoglycemia. J Diabetes Complications 2014;28(4):565-572. [doi: 10.1016/j.jdiacomp.2014.03.002] [Medline: 24685363]
4. Shafiee G, Mohajeri-Tehrani M, Pajouhi M, Larijani B. The importance of hypoglycemia in diabetic patients. J Diabetes Metab Disord 2012 Oct 01;11(1):17 [FREE Full text] [doi: 10.1186/2251-6581-11-17] [Medline: 23497433]
5. Rubin DJ, Golden SH. Hypoglycemia in non-critically ill, hospitalized patients with diabetes: evaluation, prevention, and management. Hosp Pract (1995) 2013 Feb;41(1):109-116. [doi: 10.3810/hp.2013.02.1016] [Medline: 23466973]
6. Boucai L, Southern WN, Zonszein J. Hypoglycemia-associated mortality is not drug-associated but linked to comorbidities. Am J Med 2011 Nov;124(11):1028-1035 [FREE Full text] [doi: 10.1016/j.amjmed.2011.07.011] [Medline: 22017781]
7. Hawkins K, Donihi AC, Korytkowski MT. Glycemic management in medical and surgical patients in the non-ICU setting. Curr Diab Rep 2013 Feb;13(1):96-106. [doi: 10.1007/s11892-012-0340-1] [Medline: 23100037]
8. Umpierrez GE, Pasquel FJ. Management of inpatient hyperglycemia and diabetes in older adults. Diabetes Care 2017 Apr;40(4):509-517 [FREE Full text] [doi: 10.2337/dc16-0989] [Medline: 28325798]
9. Wang Y, Eldridge N, Metersky ML, Verzier NR, Meehan TP, Pandolfi MM, et al. National trends in patient safety for four common conditions, 2005-2011. N Engl J Med 2014 Jan 23;370(4):341-351 [FREE Full text] [doi: 10.1056/NEJMsa1300991] [Medline: 24450892]
10. Amori RE, Pittas AG, Siegel RD, Kumar S, Chen JS, Karnam S, et al. Inpatient medical errors involving glucose-lowering medications and their impact on patients: review of 2,598 incidents from a voluntary electronic error-reporting database. Endocr Pract 2008;14(5):535-542. [doi: 10.4158/EP.14.5.535] [Medline: 18753094]
11. Alrwisan A, Ross J, Williams D. Medication incidents reported to an online incident reporting system. Eur J Clin Pharmacol 2011 May;67(5):527-532. [doi: 10.1007/s00228-010-0986-z] [Medline: 21240481]

12. Cobaugh DJ, Maynard G, Cooper L, Kienle PC, Vigersky R, Childers D, et al. Enhancing insulin-use safety in hospitals: practical recommendations from an ASHP Foundation expert consensus panel. Am J Health Syst Pharm 2013 Aug 15;70(16):1404-1413. [doi: 10.2146/ajhp130169] [Medline: 23903479]
13. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005 Apr 02;330(7494):765 [FREE Full text] [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]
14. Kwan JL, Lo L, Ferguson J, Goldberg H, Diaz-Martinez JP, Tomlinson G, et al. Computerised clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials. BMJ 2020 Sep 17;370:m3216 [FREE Full text] [doi: 10.1136/bmj.m3216] [Medline: 32943437]
15. Jenders RA. Advances in clinical decision support: highlights of practice and the literature 2015-2016. Yearb Med Inform 2017 Aug;26(1):125-132 [FREE Full text] [doi: 10.15265/IY-2017-012] [Medline: 29063552]
16. McRee L, Thanavaro JL, Moore K, Goldsmith M, Pasvogel A. The impact of an electronic medical record surveillance program on outcomes for patients with sepsis. Heart Lung 2014;43(6):546-549. [doi: 10.1016/j.hrtlng.2014.05.009] [Medline: 24961633]
17. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med 2015 Aug 05;7(299):299ra122. [doi: 10.1126/scitranslmed.aab3719] [Medline: 26246167]
18. Wilson F, Greenberg J. Acute kidney injury in real time: prediction, alerts, and clinical decision support. Nephron 2018;140(2):116-119 [FREE Full text] [doi: 10.1159/000492064] [Medline: 30071528]
19. Kilpatrick CR, Elliott MB, Pratt E, Schafers SJ, Blackburn MC, Heard K, et al. Prevention of inpatient hypoglycemia with a real-time informatics alert. J Hosp Med 2014 Oct;9(10):621-626. [doi: 10.1002/jhm.2221] [Medline: 24898687]
20. Meyfroidt G, Wouters P, De Becker W, Cottem D, Van den Berghe G. Impact of a computer-generated alert system on the quality of tight glycemic control. Intensive Care Med 2011 Jul;37(7):1151-1157. [doi: 10.1007/s00134-011-2159-7] [Medline: 21369814]
21. Mathioudakis NN, Everett E, Routh S, Pronovost PJ, Yeh H, Golden SH, et al. Development and validation of a prediction model for insulin-associated hypoglycemia in non-critically ill hospitalized adults. BMJ Open Diabetes Res Care 2018;6(1):e000499 [FREE Full text] [doi: 10.1136/bmjdrc-2017-000499] [Medline: 29527311]
22. Mathioudakis NN, Abusamaan MS, Shakarchi AF, Sokolinsky S, Fayzullin S, McGready J, et al. Development and validation of a machine learning model to predict near-term risk of iatrogenic hypoglycemia in hospitalized patients. JAMA Netw Open 2021 Jan 04;4(1):e2030913 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.30913] [Medline: 33416883]
23. Cruz P, Blackburn MC, Tobin GS. A systematic approach for the prevention and reduction of hypoglycemia in hospitalized patients. Curr Diab Rep 2017 Oct 05;17(11):117. [doi: 10.1007/s11892-017-0934-8] [Medline: 28980145]
24. Mathioudakis N, Jeun R, Godwin G, Perschke A, Yalamanchi S, Everett E, et al. Development and implementation of a subcutaneous insulin clinical decision support tool for hospitalized patients. J Diabetes Sci Technol 2019 May;13(3):522-532 [FREE Full text] [doi: 10.1177/1932296818798036] [Medline: 30198324]
25. Umpierrez GE, Hellman R, Korytkowski MT, Kosiborod M, Maynard GA, Montori VM, Endocrine Society. Management of hyperglycemia in hospitalized patients in non-critical care setting: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 2012 Jan;97(1):16-38. [doi: 10.1210/jc.2011-2098] [Medline: 22223765]
26. Campbell R. The five "rights" of clinical decision support. J AHIMA 2013 Oct;84(10):42-7; quiz 48. [Medline: 24245088]
27. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
28. Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R, with the HITEC Investigators. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Med Inform Decis Mak 2017 Apr 10;17(1):36-45 [FREE Full text] [doi: 10.1186/s12911-017-0430-8] [Medline: 28395667]
29. Gregory ME, Russo E, Singh H. Electronic health record alert-related workload as a predictor of burnout in primary care providers. Appl Clin Inform 2017 Jul 05;8(3):686-697 [FREE Full text] [doi: 10.4338/ACI-2017-01-RA-0003] [Medline: 28678892]
30. Hussain MI, Reynolds TL, Zheng K. Medication safety alert fatigue may be reduced via interaction design and clinical role tailoring: a systematic review. J Am Med Inform Assoc 2019 Oct 01;26(10):1141-1149 [FREE Full text] [doi: 10.1093/jamia/ocz095] [Medline: 31206159]

Abbreviations

BPA: best practice advisory
CDS: clinical decision support
CDSS: clinical decision support system
EMR: electronic medical record
ICU: intensive care unit
IM: internal medicine

NP: nurse practitioner
OB/GYN: obstetrics/gynecology

Edited by A Kushniruk; submitted 13.06.21; peer-reviewed by P Cruz, M Salimi, I Mircheva, R Marshall; comments to author 26.08.21; revised version received 08.09.21; accepted 11.09.21; published 26.11.21.

Please cite as:
Mathioudakis N, Aboabdo M, Abusamaan MS, Yuan C, Lewis Boyer L, Pilla SJ, Johnson E, Desai S, Knight A, Greene P, Golden SH
Stakeholder Perspectives on an Inpatient Hypoglycemia Informatics Alert: Mixed Methods Study
JMIR Hum Factors 2021;8(4):e31214
URL: https://humanfactors.jmir.org/2021/4/e31214
doi:10.2196/31214
PMID:34842544

Original Paper

Exploring Usability Issues of a Smartphone-Based Physician-to-Physician Teleconsultation App in an Orthopedic Clinic: Mixed Methods Study

Songphan Choemprayong[1,2], MLIS, PhD; Chris Charoenlap[3], MD; Krerk Piromsopa[4], PhD

[1]Department of Library Science, Faculty of Arts, Chulalongkorn University, Bangkok, Thailand

[2]Behavioral Research and Informatics in Social Science Research Unit, Sasin School of Management, Chulalongkorn University, Bangkok, Thailand

[3]Department of Orthopaedic, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

[4]Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand

Corresponding Author:
Chris Charoenlap, MD
Department of Orthopaedic
Faculty of Medicine
Chulalongkorn University
Rama IV Road, Pathumwan
Bangkok, 10330
Thailand
Phone: 66 081 552 4224
Fax: 66 02 256 4625
Email: chris.cha@chula.ac.th

Abstract

Background: Physician-to-physician teleconsultation has increasingly played an essential role in delivering optimum health care services, particularly in orthopedic practice. In this study, the usability of a smartphone app for teleconsultation among orthopedic specialists was investigated to explore issues informing further recommendations for improvement in the following iterations.

Objective: This study aimed to explore usability issues emerging from users' interactions with MEDIC app, a smartphone-based patient-centered physician-to-physician teleconsultation system.

Methods: Five attending physicians in the Department of Orthopedics in a large medical school in Bangkok, Thailand, were recruited and asked to perform 5 evaluation tasks, namely, group formation, patient registration, clinical data capturing, case record form creation, and teleconsultation. In addition, one expert user was recruited as the control participant. Think aloud was adopted while performing the tasks. Semistructured interviews were conducted after each task and prior to the exit. Quantitative and qualitative measures were used to identify usability issues in 7 domains based on the People At the Centre of Mobile Application Development model: effectiveness, efficiency, satisfaction, learnability, memorability, error, and cognitive load.

Results: Several measures indicate various aspects of usability of the app, including completion rates, time to completion, number of clicks, number of screens, errors, incidents where participants were unable to perform tasks, which had previously been completed, and perceived task difficulty. Major and critical usability issues based on participant feedback were rooted from the limitation of screen size and resolution. Errors in data input (eg, typing errors, miscalculation), action failures, and misinterpretation of data (ie, radiography) were the most critical and common issues found in this study. A few participants did not complete the assigned tasks mostly owing to the navigation design and misreading/misunderstanding icons. However, the novice users were quite positive that they would be able to become familiar with the app in a short period of time.

Conclusions: The usability issues in physician-to-physician teleconsultation systems in smartphones, in general, are derived from the limitations of smartphones and their operating systems. Although some recommendations were devised to handle these usability issues, usability evaluation for additional development should still be further investigated.

(JMIR Hum Factors 2021;8(4):e31130) doi:10.2196/31130

KEYWORDS

teleconsultation; remote consultation; mobile applications; usability; orthopedics; physician-to-physician consultation; electronic medical records; mobile phone

Introduction

Comprehensive personalized care has become a desirable model for health care systems in many countries. Such a model requires the integration of multiple stakeholders and systems to provide patient-centered services. In addition, the efficiency of health care systems, particularly in terms of collecting, storing, analyzing, and accessing patient records and related data is now even more critical to the quality of health care service delivery. Information and communication technologies have advanced health care services in numerous ways, particularly in reducing medical errors, paper consumption, physical storage space, and time. Electronic medical records (EMRs), for instance, have been widely adopted since they play an essential role in the data repository as well as a point of reference in communication between health care providers and patients. Although EMRs are concerned with how health care providers manage patient records, the modern health care system requires collaboration between health care providers, in the expectation of improving the quality of diagnosis and the treatment process, thereby ensuring the quality of data and increasing trust among providers as well as between providers and patients [1].

Teleconsultation is broadly defined and used to explain the remote communication between at least 2 parties in conducting the health care process and services (eg, between a primary care physician and a specialist, between a physician and a nurse, between a resident and a supervising physician, and between a physician/nurse and a patient) [1]. Owing to the advancement of information and communication technology, teleconsultation can be delivered via various channels, for instance, telephone [1,2], video conference systems [3], instant messengers [4-6], and smartphone apps [7]. In particular, the use of mobile devices has increased worldwide since the first introduction of smartphones in the late 2000s. The International Telecommunication Union [8] estimates that there were almost 8 billion mobile cellular subscriptions worldwide in 2020. In Thailand, there were in excess of approximately 119 million mobile subscriptions in 2020 [9]. Mobile phones have become a major platform, bypassing desktops and websites in many other areas.

Teleconsultation has become more common in orthopedic care [10,11], particularly in terms of telemonitoring, teleradiography, and telesurgery [12]. The COVID-19 pandemic has accelerated the adoption of telemedicine, despite criticism and resistance in certain specialties [10,13,14]. There have been several evaluation studies on the effectiveness of teleconsultation in terms of data quality and clinical outcomes (eg, length of stay, user satisfaction, economic evaluation [10,15]). Although the results from systematic review studies cannot confirm the clinical benefits of teleconsultation [15], it is apparent that orthopedic specialists, particularly surgeons, prefer teleconsultation over traditional office visits with patients [10].

Considering that telemedicine and collaborative practice have been increasingly adopted in orthopedic care, a patient-focused teleconsultation platform allows specialists and health care providers to be able to access up-to-date patient records anytime and anywhere. Where diagnosis and prescription are needed, health care providers can update patient records and provide consultation on the go. However, developing a mobile app, in general, can have numerous usability challenges. For instance, limited screen size and resolution restrict the capacity to display large-scale information. Moreover, screen size and resolution may affect the performance of data input, particularly when typing and selecting from a list [16]. Other factors that can affect the usability of mobile apps include distractions during use, connection speed, and processing power [17]. Designers and developers of mobile apps and websites may have to compromise their design in numerous ways, for example, by segmenting and presenting information in multiple pages. Although the usability of mobile apps has been widely studied [17-20], their usability in physician-to-physician teleconsultation has seldom been investigated [7,21,22]. Abundant usability evaluations in health care systems focus on EMRs in the desktop environment [23-25]. Even in the usability studies of mobile EMR systems [26-32], most of them tend to apply generic frameworks rather than those developed for mobile app or teleconsultation specifically. Kim et al [33] called for further explorations of the feasibility, particularly from a usability perspective, of mobile apps on smartphones among physicians.

Harrison et al [18] point out that most mobile usability models focus only on 3 basic attributes, that is, effectiveness, efficiency, and satisfaction, overlooking other essential attributes such as cognitive load. People At the Centre of Mobile Application Development (PACMAD), an evaluation framework, was developed and tailored to address the usability of mobile apps. Based on the International Organization of Standardization and the famous Nielsen's model [34], PACMAD covers 7 relevant usability attributes, that is, learnability, efficiency, effectiveness, errors, memorability, satisfaction, and cognitive load. Although the constructs in this model cover a wide range of usability aspects and the model has been widely used in various contexts of use, goals, and groups of users, including patient-based mobile health apps [35-37], its application in physician-to-physician consultation mobile apps is very limited. As the framework is applicable to the usability of mobile apps in general, this study adopts PACMAD as the theoretical and analytical framework.

Using a heuristic evaluation approach [38], this study aimed to explore usability issues emerging from interaction with a patient-oriented physician-to-physician teleconsultation app on a smartphone device. Although the app developed can be applied to other settings, this study uses an orthopedic clinic as a setting to control the complexity of the task and the potential confounding determinants [23]. In addition to informing the recommendations and guidelines for the design and development of medical apps on small mobile devices, this study also sheds

some light on the feasibility of teleconsultation apps among physicians on smartphones, which are more pervasive and portable. Further, this study investigates the applicability of PACMAD, a general usability framework of mobile devices, in the context of physician-to-physician teleconsultation.

Methods

Participants

Nielsen [39] argues that, in exploring usability issues among homogenous users, the first 3 users will help discover problems in an exponential manner. The data are hypothetically saturated after the fifth user. In addition, the primary users of the current version of MEDIC are orthopedic specialists and physicians who normally provide consultations with each other. As a single-site study, the study site was an orthopedic clinic in a large medical school in Bangkok, Thailand, housing around 60 orthopedic specialists and physicians. The study enrolment was announced in the department meeting and all participants joined voluntarily. One orthopedist who regularly used the MEDIC app was recruited to be the control participant. Four specialists and 1 resident who had none or a few experiences of using MEDIC were recruited in this study. To control the complexity of the task and the variability of the platform, the usability tests were conducted using the MEDIC iPhone operating system platform only. Therefore, all participants had to be current iPhone users. The study protocol (IRB 756/62) was approved by the Institutional Review Board, Faculty of Medicine, Chulalongkorn University.

Data Collection

This study was conducted in a controlled setting in the Clinical Skill and Simulation Center (Figure 1). Upon arrival, participants were introduced to the MEDIC app through a 5-minute video presentation providing information on the main features and functions of the app. Thereafter, participants were asked to complete a usability test session. The overall test for each participant took about 90 minutes to complete. There were 5 tasks given to the participants to complete individually. In the context of an orthopedic clinic, these tasks were designed to cover the basic functions of the app and the data capturing process in simulated clinical situations. The participants were required to use the app on the provided iPhone 8. They were also allowed the use of other apps on the device to complete the tasks. Owing to the collaborative nature of the teleconsultation work, the users had to create a private group serving as a sharing space to communicate between physicians. The group could be utilized for a clinical unit, a discussion about a specific case or a group of cases, a research project, or a certain task force. Thus, the first task assigned the study participants to create a group and invite other users to join the newly created group. Since MEDIC is designed to be a patient-focused teleconsultation app, MEDIC allows physicians to create and manage patient records within a group space. Only physicians in the group can view and manage patient records in the group. Therefore, the participants were asked to create a new patient record in task 2. In addition, to evaluate how physicians use the app to manage patient records in an environment similar to the natural setting, the study participants were asked to collect patient information by using the app in a simulated clinical visit in task 3.

Another key function of MEDIC is to allow physicians to communicate in a standardized manner via the data collection form. There are several scales and measures that are essential for the clinical management of patients. Some are standardized, while some are tailored and customized within a group. The data collection form function allows physicians to create and customize a form to be used among the physicians in the group. Therefore, task 4 was designed to investigate how the participants used the data collection form by asking them to create a form to collect a Mangled Extremity Severity Score, which is one of the most common scales used in orthopedic clinical practices. Although the first 4 tasks aimed to observe how the participants created and managed data within the app from a sender perspective (eg, medical students, resident physicians, referrers), the last task was designed to evaluate the usability of the app from the perspective of the receiver (eg, peers, advisors, supervisors, referees). The participants were asked to review existing patients' records who were asked for consultation. The participants were asked to provide clinical opinions as well as to enter a diagnosis and treatment plan in the consulting case's record. While performing the tasks, participants were asked to speak aloud their thoughts to the researchers. Participant activities were logged using the video camera and screen recording function of the iPhone. The description of all the tasks is shown in Textbox 1.

Although there was no limitation on the completion time, participants could leave tasks incomplete at any point in the task. The radiograph used in task 5 is shown in Figure 2. After completing each task, the participants were interviewed using a semistructured interview approach to obtain detailed information regarding their behavior and experience. Additionally, an exit interview was also conducted after all the tasks had been completed. All user actions were recorded using a screen capturing app and video recording. One of the research team members also observed the participants and recorded their actions in an observational form.

Figure 1. The camera setup for observing research participant interactions with the MEDIC app, while doing history taking and examination with the simulated patient in task 3.

Textbox 1. Tasks in the usability test of the smartphone-based physician-to-physician teleconsultation app.

Task 1: Creating a group, then adding team members and form into the group

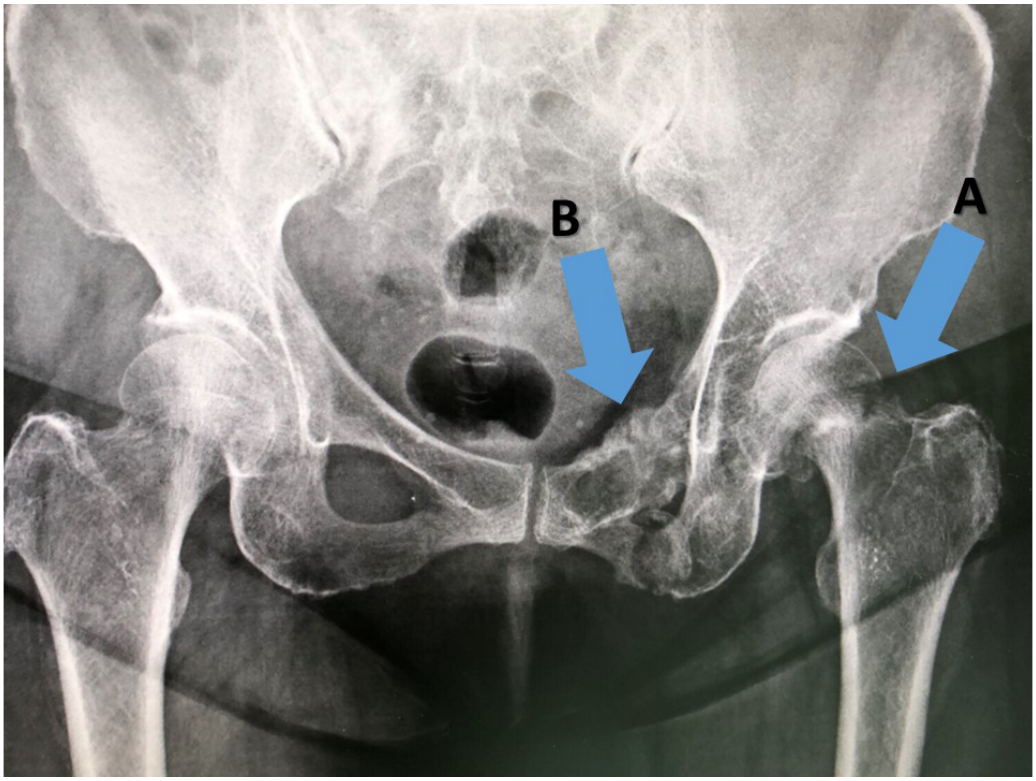
Task 2: Registering the patient into the group

Task 3: Using MEDIC app during a clinical encounter with a simulated patient, including the recording of a radiograph, a photo of the affected body part (the knee), a physical examination, diagnosis, and plan management. The simulated patient was informed about the case and trained by the researcher. The scenario for task 3 was a 42-year-old female patient presenting with chronic pain in her right knee for 3 months. A plain radiograph showed that she had osteoarthritis, which is a common degenerative condition of the knee joint. The treatment included medication, physical therapy, and surgery.

Task 4: Create a new record form, a Mangled Extremity Severity Score, which is used to assess limb-salvage potential of traumatic extremity. Prior to this task, participants were given a brief video introduction on how to create the form.

Task 5: Using MEDIC for a teleconsultation of orthopedic trauma cases. Participants were asked to review a patient's radiograph and then provide a diagnosis and treatment plan as well as clinical opinion. The task 5 case was a 35-year-old female who had had a traffic accident 2 hours before arrival at the emergency room. A plain radiograph showed a fractured neck of the left femur, which is the proximal part of the thigh bone, and pubic rami fracture of the pelvic bone.

Figure 2. Plain radiograph image, which is used in task 5, showing fractured neck of left hip (femur) (A) and pelvis (pubic rami) (B).



Data Analysis

The data from the think-aloud protocol and the semistructured interviews were transcribed. The observation notes were validated with the video recording. Quantitative data, for instance, time, number of clicks, and number of screens used, were recorded in MS Excel and analyzed using descriptive statistics by comparing with expert users' performance. Thematic analysis using a deductive coding approach was adopted to analyze the transcripts, observation notes, and screen and video recordings by using the qualitative data analysis software, NVivo version 12 (QSR International). The primary coding scheme adopted PACMAD usability attributes [18], including effectiveness, efficiency, satisfaction, learnability, memorability, error, and cognitive load. The coding was conducted by 2 assessors independently. The codes and categories were then compared. Any disagreements were resolved by discussion between the 2 assessors.

Evaluation Measures

To explore the usability issues of MEDIC, Table 1 shows the measures collected during and after tasks. To triangulate the results, the data were derived from 3 data sources: observation, think-aloud responses, and interviews.

It is apparent that certain measures were attributed to more than one usability domain, for instance, perceived task difficulty addresses both learnability and cognitive load. The time used to accomplish assigned tasks was also used to evaluate learnability and efficiency. For cognitive load, although the NASA Task Load Index is normally recommended [18,40], we considered Flood's hypothetical approach [41] instead since it specifically addresses the cognitive load in a mobile environment in the context of clinical and health care practices. Furthermore, 2 additional measures were collected to understand user characteristics: (1) familiarity with each task assigned (rating on 1-7 Likert scale) and (2) familiarity with heavy-loaded tasks (eg, writing, filling out a form) on mobile platforms.

Table 1. Measures of usability.

Usability attributes	Data sources		
	Observations during task performances	Posttask interviews	Exit interviews
Effectiveness	Completion rate using the Laplace method	—[a]	—
	Comparing the time to completion, number of clicks and number of screens with expert performance		
Efficiency	Time to completion	—	—
	Number of clicks		
	Number of screens used		
Satisfaction	—	—	Perceived potential impact of the app on the effectiveness of current workflow
Learnability	Time to completion	Perceived task difficulty	—
Memorability	—	Number of incidents where participants were unable to perform a task, which had previously been completed	—
Error	Incidents when errors occurred	—	—
Cognitive load	Distractions during task performance	Perceived task difficulty	—

[a]Not available.

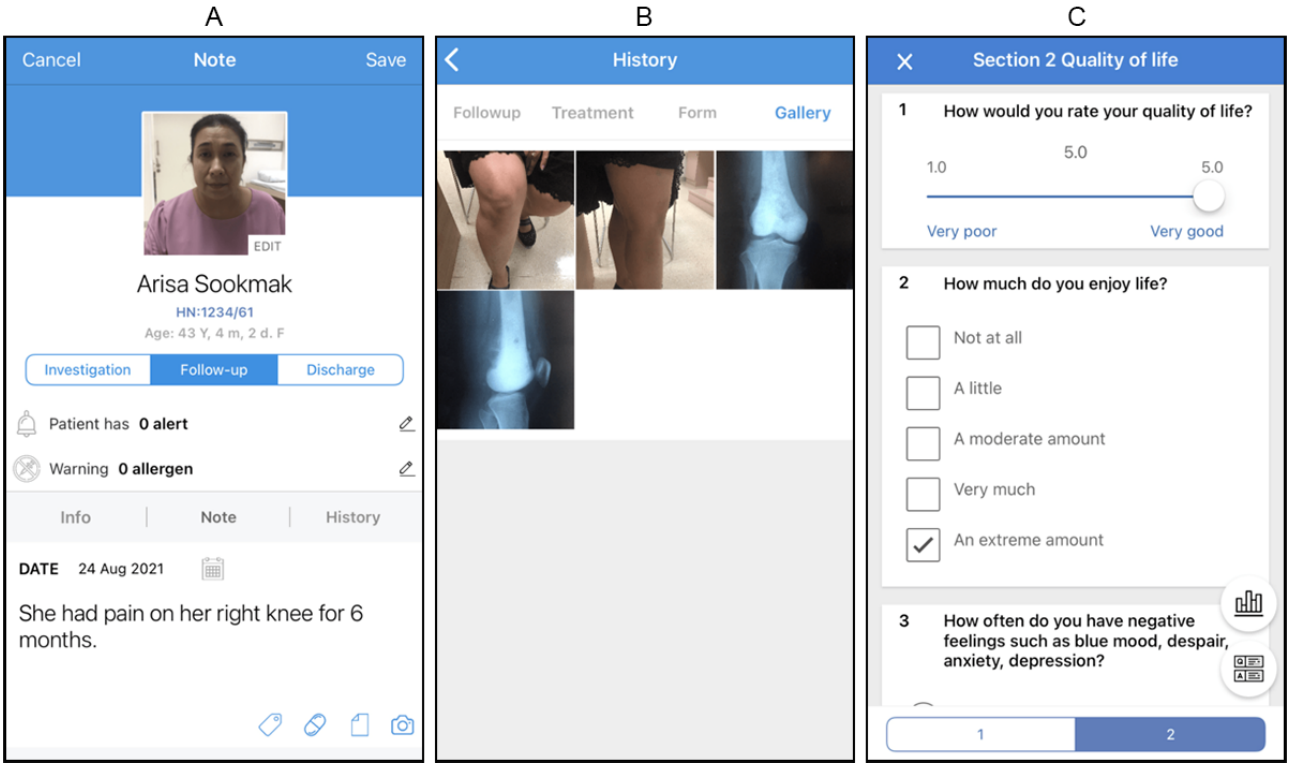
MEDIC: Smartphone Physician-to-Physician Teleconsultation App

MEDIC, developed by Deverhood, Thailand, is a smartphone teleconsultation app for physicians to communicate with each other in various settings. As the app is designed to support patient-centered care, the main features of the app include patient medical data such as medical history, physical examination, clinical images, and diagnostic questionnaires. Health care providers can access, collect, and modify data from both desktop and mobile platforms, including both iPhone operating system and Android. However, the testing version in this study was on the iPhone operating system platform to control the environment. All data were to be uploaded to the cloud server; therefore, an internet connection was required while using the app. Designed to support collaboration among physicians and specialists, the MEDIC interface is divided to support 4 main tasks, namely, forming a team, data form creation, data recording, and data reviewing. The first task begins with forming a team such as a research group or a multicenter collaboration by creating a group and adding members. To invite members to the group, all teammates must have accounts with MEDIC. Group members can be removed or included by the group administrator.

Although MEDIC has been designed to collect generic patient records (eg, demographic, diagnosis, medical history, treatment), the app also allows physicians to create a data form to support their specialty, such as a case record form and functional score. However, licensed questionnaires should be authorized by the licensed owner in advance. There are 10 data input formats that can be used in the form: check boxes, drop-down lists, multiple choice, linear scale, multiple choice grids, free text, number text, date-time, picture, and video link. The content of the form can be organized into a section. Each question can be set as a required status, which should be completed or the form cannot be submitted. Relevant forms should be assigned to the related group. The data capturing process begins with patient registration with the group by entering a general profile and collecting data using a free text box, form, and camera tool. All recorded data can be reviewed by pressing the previous history tab, which shows free text history, forms, and images that were recorded in the past. Apart from all the main features, this app allows users to fill out their profiles for reference and a setup passcode lock to increase data security. The main features of the app are shown in Figure 3.

Apart from the capability of clinical data capturing, MEDIC is suitable for teleconsultation between health care personnel. The MEDIC app provides organized information, including history, laboratory findings, and clinical and radiological images, and patient condition and management are automatically sorted in a chronological order so that it is convenient for reviewing disease progression and treatment plan. Patient data privacy protection is improved by using MEDIC instead of social networks such as WhatsApp or Facebook messenger because the data access is limited to authorized persons for use in patient management.

Figure 3. Three screenshots of the MEDIC app: patient history note (A), image gallery (B), and record form (C).



Results

Participants' Characteristics

Demographics

All participants were males whose age ranged from the late 20s to early 40s. One of the participants was in the final year of orthopedic resident training while the others were board-certified orthopedists. All participants, reportedly, were highly familiar with and had been using smartphones for many years. One expert was an orthopedist who used the MEDIC app regularly and had been involved with the development of the app.

Familiarity With Usability Tasks

The participants were asked to declare how familiar they were with each given task. As shown in Table 2, they tended to be the most familiar with task 2 and task 4 (median=6). For the other tasks, familiarity was distributed among the 5 participants.

We also asked the participants to provide feedback on their familiarity with heavy-loaded tasks on mobile platforms. All participants said they were familiar with multitasking on smartphones with regard to work-related tasks.

Table 2. Familiarity with tasks (N=5).

Task	Median (range)
1	3 (1-7)
2	6 (4-7)
3	3 (1-5.5)
4	6 (4-6.5)
5	4 (1-7)

Usability of the MEDIC App

The following section reports different measures covering all 7 usability dimensions of PACMAD in the smartphone-based physician-to-physician teleconsultation app, that is, completion rates, time to completion, number of clicks, number of screens, errors, incidents where the participant was unable to perform a task, which had previously been completed, and perceived task difficulty.

Completion Rates

As shown in Table 3, all participants completed tasks 2, 3, and 4 (completion rate=100%; Laplace=0.86). One participant did not complete task 1 and 5 (completion rate=80%; Laplace=0.71). One participant did not complete task 1 because he failed to add a user (another physician) into the created group. For task 5, the participant could not locate or contact another physician for consultation for a specific case.

Table 3. Completion rates (N=5).

Task	Completion rate, n (%)	Laplace
1	4 (80)	0.7143
2	5 (100)	0.8571
3	5 (100)	0.8571
4	5 (100)	0.8571
5	4 (80)	0.7143

Time to Completion (Minutes)

As shown in Table 4, the participants completed task 1 taking about 3 times longer (median=6 minutes) than the time used by the expert (median=2 minutes). For task 2, the median time used by participants (median=2 minutes) was about the same as the time used by the expert. For task 3, the participants used about 16 minutes to complete the task, while the expert used about 6 minutes. It is noteworthy that task 3 involved interviewing a simulated patient. Therefore, the range of time used was from 8 minutes to 24 minutes. The participants completed task 4 using about 8 minutes, approximately 1.25 times more than the time used by the expert. The median time spent by the participants was approximately 2 times more than that spent by one of the experts (6 minutes and 3 minutes, respectively).

Table 4. Time (minutes) used by given tasks (N=5).

Task	Median (range)	Expert
1	6 (2-8)	2
2	2 (1-6)	2
3	16 (8-24)	6
4	8 (6-11)	6
5	6 (4-8)	3

Number of Clicks

We also observed the number of mouse clicks during each task, as illustrated in Table 5. For all tasks, except task 2, the median number of clicks by the participants was higher than those by the expert (41 versus 27 in task 1, 88 versus 27 in task 3, 92 versus 68 in task 4, and 36 versus 18 in task 5). The greatest difference between the median number of clicks by participants and the number of clicks by the expert was in task 3 (about 3.26 times higher). However, there was 1 participant who could complete task 3 within 17 clicks, which was lower than the number of clicks by the expert. For task 2, all participants completed the task by using a lower number of clicks than that used by the expert (13 and 19, respectively).

Table 5. Number of clicks by given tasks (N=5).

Task	Median (range)	Expert
1	41 (15-41)	27
2	13 (10-14)	19
3	88 (17-133)	27
4	92 (74-131)	68
5	36 (28-45)	18

Number of Screens Used

In terms of screens used, as shown in Table 6, the median number of screens used in all tasks was higher than that of the screens used by the expert. For task 1, the median number of screens used was 25 screens, while the expert used 17 screens. The participants completed task 2 by using 6 screens (median) compared to 4 screens by the expert. The median number of screens used by the participants in task 3 was much higher than that used by the expert (46 and 17, respectively). In task 4, the median number of screens used by the participants was 23 screens, while the expert used only 17 screens to complete this task. The participants completed task 5 by using 31 screens (median) compared to 12 screens used by the expert. It is also noteworthy that some participants used fewer screens to complete tasks 1, 3, and 4, compared to the number of screens used by the expert.

Table 6. Number of screens visited by given tasks (N=5).

Task	Median (range)	Expert
1	25 (10-40)	17
2	6 (4-7)	4
3	46 (9-72)	17
4	23 (14-29)	17
5	31 (16-38)	12

Errors

In this study, we observed errors through direct observation as well as through user feedback. Three types of errors were identified: action failure, data inaccuracy, and error recovery failure.

The first type of error was action failure. The participants could not complete certain activities. For example, in task 3, one participant entered the diagnosis information into a wrong section. It is apparent that the term "form" was used in multiple sections where it was meant differently depending on in which section it appeared. Therefore, this discrepancy led to subsequent confusion and data entry error. Another major action failure dealt with navigation issues. Some participants could not navigate the app or correctly locate the section where they expected to complete tasks. Apparently, they did not understand the vocabulary and icons used. In addition, in task 1, where they were asked to create a group and add a member into the group, most participants took a lot of time navigating through the app to add a member into a created group. Most of them used a trial-and-error approach by browsing and clicking all buttons to see if they were helpful.

The action failure was found to be related to data inaccuracy, which is another type of error found in this study. In task 4, where they were asked to create and complete an assessment form, 3 participants entered data into incorrect fields because they were confused about how to create and fill out the information in the form. In addition, 1 participant incorrectly input the patient birth year from 1978 to 1987. This incident was led by the discrepancy between the calendar year system required in the app (Gregorian calendar year) and the official local calendar system (Buddhist calendar year). The participant had to manually add an extra step to convert the difference between the two calendar year systems, which increases the risk of data inaccuracy.

Another error was the frustration to recover after encountering an error. There were a few incidents where the participants found certain mistakes where they would have liked to make some changes. However, they could not find a solution to this for 2 main reasons. First, the app did not have an edit function for specific tasks. Second, the participants' mental model about how to edit did not match the edit operation in the app. For instance, participant 2 tried to create a form to collect patient data in task 4. However, after adding a question, the participant found that he had misplaced the order of the question. He struggled to find a workaround to reorder the question. He ended up deleting the entire questionnaire and started creating a new form instead.

Incidents Where a Participant Was Unable to Perform a Task That Had Previously Been Completed

Based on a direct observation during the tests and validated by 2 observers, there was no incident where participants were unable to perform a task, which had previously been completed.

Perceived Task Difficulty

To evaluate the cognitive load during task performance, we asked the participants to rate perceived task difficulty on a 7-point Likert scale (1, not difficult at all; 7, most difficult). Based on the medians, as shown in Table 7, the perceived difficulties in all tasks were considered as moderate to most difficult depending upon each task assignment. However, considering the range, some participants rated tasks 1, 3, and 5 as less difficult.

Table 7. Perceived task difficulty rating (N=5).

Task	Median (range)
1	3 (1-7)
2	6 (4-7)
3	3 (1-7)
4	6 (4-5.5)
5	4 (1-5.5)

Moreover, we asked the participants to provide feedback supporting their ratings. We analyzed their feedback in relation to cognitive load. For those who rated tasks as less difficult, the rationale supporting their perception included the clarity of the interface, compatibility with their workflow, and the familiarity of the task to current practices (eg, creating a group and adding a group member). In addition to the perceived task difficulty, we also observed the distractions during task completion in all tasks. However, those who perceived given tasks as difficult provided feedback that certain actions required additional

resources (eg, time, memory). Participant 2 commented on the difficulty of converting the date of birth from the Buddhist calendar year to the Gregorian calendar.

> *…Well, for entering the date [of birth], sometimes, I need time to think and to fill in the information. It takes time to do so. But we really don't have a lot of time for each patient.* [Participant 2]

For some participants, typing on a small screen was another intensive task. Participants reported some difficulty in typing on the screen. One participant said that his fingers were too big for the screen, causing misalignment with the keyboard. Participant 3 compared typing with taking a photo of a handwriting chart. He felt that writing on paper and taking pictures was faster than typing. Especially when typing while meeting patients, the participants felt that they were distracted by how much attention was required to focus on what was typed rather than the interaction with the patient. To compromise the cognitive load during typing, participant 3 decided to keep the notes concise, for instance, by using abbreviations and short phrases. It is noteworthy that omitting certain information in the medical record can cause cognitive load in recalling the information. For example, time-sensitive information requires a specific unit. Participant 2 expressed his frustration regarding the ambiguity caused by lack of contextual information.

> *…For the question 'how long ago did the patient have surgery? I usually put the unit, like month, in the chart. If it only has a number, I have no idea what this number means. For example, if I see number 6, what does it mean? 6 months? 6 days? or 6 years? I cannot tell.... It makes it a bit difficult to communicate with the patient.* [Participant 2]

Another incident concerned unfamiliarity with the form creation process in task 5. MEDIC allows physicians to create a customized form to collect certain information for further evaluation. For those who have never created a form before, they were confused by the terms used in the app, for example, section, question, dropdown menu, and check box. They spent a great amount of time trying to figure out how to create a form.

Another task that was rated difficult in some responses was form creation in task 4. A number of participants were unfamiliar with the vocabulary and the process of form creation. For example, the app allows the users to separate a questionnaire into multiple sections. However, the assigned scale, Mangled Extremity Severity Score, contains 4 questions, which do not require subsections. A couple of participants were confused about the term "section" in the form creation function. They took some time to understand the difference between the term section and question. This is perhaps partly because the items of the assigned scale are not presented in a question statement but rather in a heading format (eg, skeletal/soft-tissue injury, limb ischemia, shock, age), which could be assimilated with section titles, rather than questions.

A camera can be a useful function to capture patient records. MEDIC also provides an in-app camera function so that the user can embed photos related to patient records. The in-app camera can save time and cognitive load. However, during the test of task 3, one participant used a mobile camera instead of the in-app camera. Although the participant was able to finish the task assignment, it took him extra time and effort to completely locate the photos as well as upload them into the app.

Potential Impact of Using the App on the Current Workflow

To evaluate user satisfaction with the app, we asked all participants their opinion about the potential impact of using the app on the current workflow. All participants tended to have a positive attitude toward the impact of using the app in their practice. They thought it would likely improve the efficiency of the current workflow. The app was preferred to using instant messaging apps to communicate among physicians regarding a patient's prognosis and treatment plan.

> *...I think it is convenient to use for consultation across departments, especially for cases that need long-term care and continual discussions. It seemed impractical to use [an instant messenger app]. The problem is the chat room contains discussion and records of multiple cases. So, we have to find relevant information from a very long conversation. It is much better to get a whole patient record at once.* [Participant 1]

Nevertheless, a few participants (n=3) addressed the point that familiarity with the app would be the most substantial condition that affects the efficiency of the workflow. In addition, the performance of the app was also another factor raised by a couple of participants.

> *...Firstly, I think it's about familiarity with the tool. If you don't use it every day, you won't get used to it. Secondly, it depends on the app itself. The issues related to the app, for example, misalignment of the interface and delayed or frozen app, forced me to quit and/or restart it. Anyway, the point is that if you aren't familiar with it, it will always be difficult to use.* [Participant 2]

Discussion

Feasibility of the MEDIC App Adoption

This study aimed to discover usability issues related to the smartphone-based physician-to-physician teleconsultation system, MEDIC, by using a mixed methods approach. From the summative evaluation perspective, MEDIC seems to be promisingly satisfactory as an app providing opportunities for efficient communication among health care providers, improvement of privacy protection, and increasing accessibility to patient records outside the clinic as well as in the context of long-term care. In addition, MEDIC is designed and perceived as a patient-centric platform where all related records and documents are organized for each patient. The participants generally compared MEDIC with their current practices in recording patients records and consulting with other health care providers (ie, paper chart, EMR system, and instant messengers). Using instant messengers can be frustrating when discussing multiple issues and cases in one channel. One participant reported that he normally writes patient records on paper and

takes photos of the paper using his own smartphone. Although it can be convenient for capturing data, the quality of the photos of the records can be poor owing to the capturing process. In addition, it takes some time to retrieve the photos since such a process greatly relies on recall and memory. The current EMR system at the clinic is not portable and is not flexible in terms of serving specific needs and medical practice.

Usability Issues Related to Mismatched Mental Models

The results of participants' performance (ie, number of screens used, number of clicks, time used) in relation to the experts' performance can be considered as acceptable, considering that the majority of them had never used the app before. For memorability, the results yield a positive perspective since none of the participants forgot any actions that they had completed earlier during the test. However, usability issues were the most visible in learnability, errors, and cognitive load. Qualitative data were very helpful to explore users' mental models in addressing these issues. Although these issues emerged from the interaction with MEDIC, most of the feedback can be applied to smartphone-based teleconsultation systems in general. The most frustrating task for some participants was creating a patient evaluation form (task 4) based on the number of clicks and time spent as well as the comments during and after completing the task. Although all of them were familiar with the assigned scale, that is, Mangled Extremity Severity Score, the participants who were not familiar with the vocabulary related to electronic form and questionnaire development (eg, dropdown, checkbox, select option) found it difficult to understand the interface for the first time. They took a longer time to complete it since they applied a trial-and-error approach to become familiar with the form creation process. Nevertheless, they commented that it would take them only a couple of hours to get familiar with this task. Creating a usable data collection form (eg, measurement scales, questionnaires) has been reportedly one of the most challenging tasks in system design and development from a broad perspective [42,43]. In a clinical context, numerous established measurements have been extensively used to assist the delivery of health care services. Creating a data collection form in a mobile app based on existing paper-based questionnaires can be challenging for novice users since it may require a different mental model [44]. Users may need to be familiar with the available features, icons, and labels to effectively create a form. To address this issue, an introductory guide or a tutorial video about form creation could be useful for users who are using this function for the first time (Multimedia Appendix 1). At the same time, further studies should be conducted to investigate the mental models of novice users particularly on creating an electronic data collection form in health care settings.

Usability Issues Related to the Screen Size of Smartphones

Another common error among participants was related to the limited image resolution of the smartphone. Neither computer screens nor mobile phones were initially designed to be medical devices. The display size affects the usability performance in multiple ways [45], particularly issues related to data presentation and input. In task 5, we used the image of a pelvic fracture with osteoporosis in a consultation task. It is noteworthy that 4 out of 5 orthopedists failed to recognize pelvic fracture on the screen, even though all participants increased the magnification of the image using the zoom-in function. After this revelation, they commented that they had not seen the fracture or had not paid attention to it because the image was too small. This may reflect the observation that the typical size of mobile phone display may not provide sufficient detail for a radiographic diagnosis [37,46]. In the acute management of multiple bone fractures, underlying conditions such as osteoporosis and other metabolic bone disease should be investigated preoperatively for proper surgical preparation and medical management to prevent unexpected complications. However, Hasselberg et al [15] found that diagnosis validity does not only depend on technology but also the users' experience and physical ability. Certain solutions were suggested to mediate this issue, for instance, using an integrated DICOM viewer to provide better contrast, an ability to project onto a larger screen, and showing an image scale and other contextual information to raise awareness.

Another related issue as a consequence of screen size limitation is the misinterpretation and confusion about the image icons, which have also been addressed by other studies [7]. Owing to the limitation of screen size and the large amount of data, the design team decided to use image icons in place of text labels extensively in the app, especially for buttons. The decision led to issues related to naturalness, lack of information, and misleading information that are commonly found in other usability studies on smartphone-based health apps [37]. Participants who were using the app for the first time indicated that some of the icons were ambiguous and led to frustration. Some commented that some icons were too similar. When they were not sure what these icons were, they normally clicked to see where the buttons led to. This would cause frustration and be time consuming if the buttons did not lead to where they expected. However, some participants were successful in identifying the icons, referring to contextual elements, such as location of the icons, displaying content, and nearby icons. While replacing all icons with text labels would be immensely challenging, one possible solution is to use hover text, a tooltip text appearing when a user moves the cursor over a button. However, hover text is still not common in a number of developing platforms on touch screen devices. Other recommendations to improve the understandability of icons include removing unrelated or "unnecessary" icons, redesigning the icons to improve the distinctiveness among them, and providing a tutorial guide for the first-time user. These solutions are also suggested in other literature to avoid feature fatigue [47-49].

In addition to data presentation, the low resolution of smartphone screens can lead to data input errors [50-52]. We found that typing and clicking mistakes were omnipresent in all tasks. Participants commented that they usually had typing issues on their smartphones regardless of the app. They commented that the on-screen keyboard is too small. It is important to note that none of the participants used the swipe type function, where a user can glide his/her finger between characters. During task 3, one participant put the smartphone down and jotted down all the information on paper while talking

to the simulated patient. He commented that typing on the mobile phone screen was difficult and required a lot of attention. Typing would significantly distract him from having a conversation with the patient. Alternative input methods were also suggested by the participants to remedy this issue, for example, adding an audio recording function and speech recognition ability as well as providing contextualized word suggestions. In addition, conditional formulae, such as deactivating a button/input when it is irrelevant, would help reduce errors in typing and other calculation tasks.

Drawing was also another input method recommended by the participants. Although MEDIC allows users to take photos, some participants commented that taking pictures alone might not be enough to capture all the information they would like to add. During the consultation in task 3, one participant mentioned that he wished to annotate the images taken by either drawing or typing next to the area of interest (eg, pain site). In addition, some of the gold-standard diagnostic scales require drawing as an input. For example, the Montreal Cognitive Assessment uses clock drawing to evaluate visuoconstruction skill [53]. Therefore, the next iteration of the app development has added a drawing function as well as image annotation as shown in Figure 4. However, it is important to further investigate the usability of drawing functions on smartphone screens since it has been reported elsewhere regarding user frustration [7].

Figure 4. A prototype interface for clock drawing in Montreal Cognitive Assessment on MEDIC app.

Applicability of PACMAD Framework for Usability Evaluation of Smartphone-Based Telemedicine Apps

In this study, the PACMAD framework was applied to guide the usability evaluation of a physician-to-physician teleconsultation app. We believe this is the first study utilizing this framework in this telemedicine context. PACMAD seems to be helpful to explore a broad range of usability issues of mobile apps in line with the heuristic evaluation approach. There is currently no specific usability framework specific to smartphone-based telemedicine apps. Although PACMAD was developed as an evaluation framework tailored to the usability of mobile apps in general [18], medical apps and systems can be more sensitive in certain usability aspects, for example, naturalness, consistency, and error prevention [54]. In addition, the evaluation in medical context should consider the complexity of usability from various perspectives, including user skills, task complexity, data sensitivity, and complicated functionality. Smelcer et al [23] argued that understanding the depth and breadth of user knowledge in context is central to the usability of EMR. There are specific frameworks/guidelines addressing the usability of medical systems, in particular [55,56], such as the TURF (task, user, representation, function) framework [25,40]. Applying these frameworks may have helped explore the complexity of usability issues in medical context; however, these frameworks were developed based on the context of desktop-oriented or web-based EMRs. Some are designed based on systems interacting with patients (eg, personal health records, physician-to-patient teleconsultation). It is apparent that PACMAD is more approachable and applicable to formative evaluation approaches where researchers are less restrained from sophisticated constructs and complicated research design. Even though applying more than one framework may be achievable, we found that both frameworks are not totally compatible. For instance, TURF includes both methodological and constructive guidelines, while PACMAD focuses more on usability dimensions. It would be ideal to develop a particular framework to evaluate the usability of medical apps on a mobile (and perhaps tablet) platform.

Study Limitations and Recommendations for Future Studies

The main objective of this study was to explore the usability issues of MEDIC. Although Nielsen [39] suggests that 5 users would be sufficient to obtain the majority of issues, we found that a larger number of participants would enhance and increase

the reliability of the results, particularly in the quantitative analysis. In addition, a more heterogeneous sample would be sufficient to investigate the variation of usability of EMRs based on different user characteristics found in other studies (for instance, between attending and resident physician [57]). In addition, although MEDIC is designed as a physician-to-physician teleconsultation app, the specifications and configurations of the app are still at the early stage. Furthermore, the tasks were developed in the context of an orthopedic clinic in a large medical school. Therefore, the results from this study may not be generalizable to a larger population. MEDIC is not designed to be used as a comprehensive standalone app but to be used in conjunction with other modes of communication among physicians (eg, instant messaging, EMR systems). The feasibility of an integration between patient-based teleconsultation apps and instant messengers should be further explored with respect to patient privacy and safety. Since it seems that screen size plays an essential role in the usability issues of smartphone apps, further studies should investigate alternative measures to prevent errors in data entry, which is the most visible and concerning issue found in this study. Although the display technology of smartphones has been progressively improved, other modes of data entry such as voice, drawing, and click-and-point may be considered in terms of feasibility and usability.

Conclusion

Since this study investigates the usability of smartphone-based teleconsultation in the early stage of the iterative design process, the purpose and approach of this usability study is rather exploratory than conclusive. While applying a mixed-method approach to gain a comprehensive perspective across all usability dimensions, based on PACMAD, we found that the qualitative data provided insightful perspectives and helped us discover usability issues in numerous aspects. In addition, since the goal of this study was to explore usability issues and the population was quite homogeneous, the feedback from 5 participants was sufficient to discover usability issues in all usability dimensions. Although there are a number of opportunities to improve communication among health care team members as well as between health care professionals and patients, we found that the usability issues of smartphone-based teleconsultation platforms in this study were mostly concerned with learnability, errors, and cognitive load. We found serious issues regarding errors particularly due to the limitation of screen size and resolution. Such limitations impact on how physicians enter and view patient's records, which subsequently affect the diagnosis and treatment. Although the limitation of screen size has already been discussed in the literature, this study provides empirical evidence from a practical and user-oriented perspective. As in any early stage of development, there are still numerous opportunities for improvement, particularly regarding usability. An iterative process is planned to be adopted to develop the app to be more usable and expandable to a broader user group.

Acknowledgments

We would like to thank Praew Pinitphol, Maleepan Phasupong, and Chutimon Sookniran for assistance during data collection and manuscript preparation. The research grant has been provided by the Ratchadapiseksomphot Endowment Fund, Chulalongkorn University. The study was also partially supported by Behavior Research and Informatics in Social Science, Sasin Business School, Chulalongkorn University.

Conflicts of Interest

MEDIC was developed and financially supported by Deverhood, a medical technology company cofounded by KP and CC.

Multimedia Appendix 1
Tutorial video presentation of the MEDIC app.
[PPTX File , 25350 KB - humanfactors_v8i4e31130_app1.pptx]

References

1. Deldar K, Bahaadinbeigy K, Tara SM. Teleconsultation and clinical decision making: A systematic review. Acta Inform Med 2016 Jul 16;24(4):286-292 [FREE Full text] [doi: 10.5455/aim.2016.24.286-292] [Medline: 27708494]
2. Tian PGJ, Harris JR, Seikaly H, Chambers T, Alvarado S, Eurich D. Characteristics and outcomes of physician-to-physician telephone consultation programs: Environmental scan. JMIR Form Res 2021 Feb 23;5(2):e17672 [FREE Full text] [doi: 10.2196/17672] [Medline: 33620325]
3. Tachakra S, Hollingdale J, Uche CU. Evaluation of telemedical orthopaedic specialty support to a minor accident and treatment service. J Telemed Telecare 2001;7(1):27-31. [doi: 10.1258/1357633011936101] [Medline: 11265935]
4. Stahl I, Katsman A, Zaidman M, Keshet D, Sigal A, Eidelman M. Reliability of smartphone-based instant messaging application for diagnosis, classification, and decision-making in pediatric orthopedic trauma. Pediatr Emerg Care 2019 Jun;35(6):403-406. [doi: 10.1097/PEC.0000000000001211] [Medline: 28697157]

5. Giordano V, Koch HA, Mendes CH, Bergamin A, de Souza FS, do Amaral NP. WhatsApp Messenger is useful and reproducible in the assessment of tibial plateau fractures: Inter- and intra-observer agreement study. Int J Med Inform 2015 Feb;84(2):141-148. [doi: 10.1016/j.ijmedinf.2014.11.002] [Medline: 25468642]
6. Gulacti U, Lok U, Hatipoglu S, Polat H. An analysis of WhatsApp usage for communication between consulting and emergency physicians. J Med Syst 2016 Jun;40(6):130. [doi: 10.1007/s10916-016-0483-8] [Medline: 27083574]
7. Klingberg A, Wallis LA, Hasselberg M, Yen P, Fritzell SC. Teleconsultation using mobile phones for diagnosis and acute care of burn injuries among emergency physicians: Mixed-methods study. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e11076 [FREE Full text] [doi: 10.2196/11076] [Medline: 30341047]
8. International Telecommunication Union. World telecommunication/ICT indicators database. World Bank. 2021 Oct 29. URL: https://databank.worldbank.org/reports.aspx?dsid=2&series=IT.CEL.SETS [accessed 2021-11-21]
9. Manakitsomboon H. Number of mobile subscribers Thailand 2011-2020. Statista. 2021. URL: https://www.statista.com/statistics/995467/thailand-number-mobile-subscribers/ [accessed 2021-11-21]
10. Melian C, Kieser D, Frampton C, Wyatt MC. Teleconsultation in orthopaedic surgery: A systematic review and meta-analysis of patient and physician experiences. J Telemed Telecare 2020 Sep 01:1357633X20950995. [doi: 10.1177/1357633X20950995] [Medline: 32873138]
11. Chaudhry H, Nadeem S, Mundi R. How satisfied are patients and surgeons with telemedicine in orthopaedic care during the COVID-19 pandemic? A systematic review and meta-analysis. Clin Orthop Relat Res 2021 Jan 01;479(1):47-56. [doi: 10.1097/CORR.0000000000001494] [Medline: 33009231]
12. Behmanesh A, Sadoughi F, Mazhar FN, Joghataei MT, Yazdani S. Tele-orthopaedics: A systematic mapping study. J Telemed Telecare 2020 May 11:1357633X20919308. [doi: 10.1177/1357633X20919308] [Medline: 32393139]
13. Moisan P, Barimani B, Antoniou J. Orthopedic surgery and telemedicine in times of covid-19 and beyond: A review. Curr Rev Musculoskelet Med 2021 Apr;14(2):155-159 [FREE Full text] [doi: 10.1007/s12178-021-09693-9] [Medline: 33460020]
14. Hurley ET, Haskel JD, Bloom DA, Gonzalez-Lomas G, Jazrawi LM, Bosco JA, et al. The use and acceptance of telemedicine in orthopedic surgery during the COVID-19 pandemic. Telemed J E Health 2021 Jun;27(6):657-662. [doi: 10.1089/tmj.2020.0255] [Medline: 32931363]
15. Hasselberg M, Beer N, Blom L, Wallis LA, Laflamme L. Image-based medical expert teleconsultation in acute care of injuries. A systematic review of effects on information accuracy, diagnostic validity, clinical outcome, and user satisfaction. PLoS One 2014;9(6):e98539 [FREE Full text] [doi: 10.1371/journal.pone.0098539] [Medline: 24887257]
16. Spriestersbach A, Springer T. Quality attributes in mobile web app development. In: Bomarius F, Iida H, editors. Product Focused Software Process Improvement. PROFES 2004. Berlin: Springer; 2004:120-130.
17. Zhang D, Adipat B. Challenges, methodologies, and issues in the usability testing of mobile applications. Int J Hum Comput Interact 2005 Jul;18(3):293-308. [doi: 10.1207/s15327590ijhc1803_3]
18. Harrison R, Flood D, Duce D. Usability of mobile applications: Literature review and rationale for a new usability model. J Interact Sci 2013;1(1):1. [doi: 10.1186/2194-0827-1-1]
19. Weichbroth P. Usability of mobile applications: A systematic literature study. IEEE Access 2020 Mar 20;8:55563-55577. [doi: 10.1109/access.2020.2981892]
20. Komninos A. An introduction to usability. Interaction Design Foundation. 2020. URL: https://www.interaction-design.org/literature/article/an-introduction-to-usability [accessed 2021-11-21]
21. Boissin C, Blom L, Wallis L, Laflamme L. Image-based teleconsultation using smartphones or tablets: Qualitative assessment of medical experts. Emerg Med J 2017 Feb;34(2):95-99 [FREE Full text] [doi: 10.1136/emermed-2015-205258] [Medline: 27707791]
22. Ismirle J, Klautke H, Swierenga S, O'Connell L. Teleconsultation process for physicians working with ASD patients: insights from a usability evaluation. In: Marcus A, Wang W, editors. Design, User Experience, and Usability: Understanding Users and Contexts. Cham: Springer International Publishing; 2017:632-644.
23. Smelcer JB, Miller-Jacobs H, Kantrovich L. Usability of electronic medical records. Journal of Usability Studies. 2009. URL: https://uxpajournal.org/usability-of-electronic-medical-records/ [accessed 2021-11-25]
24. Zahabi M, Kaber DB, Swangnetr M. Usability and safety in electronic medical records interface design: A review of recent literature and guideline formulation. Hum Factors 2015 Aug;57(5):805-834. [doi: 10.1177/0018720815576827] [Medline: 25850118]
25. Zhang J, Walji MF. TURF: Toward a unified framework of EHR usability. J Biomed Inform 2011 Dec;44(6):1056-1067 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.005] [Medline: 21867774]
26. Wu RC, Straus SE. Evidence for handheld electronic medical records in improving care: A systematic review. BMC Med Inform Decis Mak 2006 Jun 20;6:26 [FREE Full text] [doi: 10.1186/1472-6947-6-26] [Medline: 16787539]
27. Wu RC, Orr MS, Chignell M, Straus SE. Usability of a mobile electronic medical record prototype: a verbal protocol analysis. Inform Health Soc Care 2008 Jun;33(2):139-149. [doi: 10.1080/17538150802127223] [Medline: 18604757]
28. Croon R, Klerkx J, Duval E. Designing a useful and usable mobile EMR application through a participatory design methodology: A case study. Los Alamitos: IEEE; 2014 Presented at: 2014 IEEE International Conference on Healthcare Informatics; September 15-17, 2014; Verona, Italy p. 176-185. [doi: 10.1109/ichi.2014.31]

29. Landman A, Emani S, Carlile N, Rosenthal DI, Semakov S, Pallin DJ, et al. A mobile app for securely capturing and transferring clinical images to the electronic health record: Description and preliminary usability study. JMIR Mhealth Uhealth 2015 Jan 02;3(1):e1 [FREE Full text] [doi: 10.2196/mhealth.3481] [Medline: 25565678]
30. Sarkar U, Gourley GI, Lyles CR, Tieu L, Clarity C, Newmark L, et al. Usability of commercially available mobile applications for diverse patients. J Gen Intern Med 2016 Dec;31(12):1417-1426 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
31. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: A usability study. JMIR Mhealth Uhealth 2014 Mar 14;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
32. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
33. Kim M, Ben RA, Clarke M, Markey M, Lage K, Aro M, et al. Usability of mobile applications supporting training in diagnostic decision-making by radiologists. In: Duffy V, editor. Digital Human Modeling. Applications in Health, Safety, Ergonomics and Risk Management: Ergonomics and Health. Cham: Springer Verlag; 2015:448-454.
34. Nielsen J. Usability Engineering. San Francisco: Morgan Kaufmann; 1994.
35. Alturki R, Gay V. Usability testing of fitness mobile app: Methodology and quantitative results. 2017 Jul Presented at: 7th International Conference on Computer Science, Engineering & Applications; September 23-24, 2017; Copenhegan, Denmark p. 97-114 URL: https://airccj.org/CSCP/vol7/csit77308.pdf [doi: 10.5121/csit.2017.71108]
36. Stiles-Shields C, Montague E, Lattie EG, Schueller SM, Kwasny MJ, Mohr DC. Exploring user learnability and learning performance in an app for depression: Usability study. JMIR Hum Factors 2017 Aug 11;4(3):e18 [FREE Full text] [doi: 10.2196/humanfactors.7951] [Medline: 28801301]
37. Kim MS, Aro MR, Lage KJ, Ingalls KL, Sindhwani V, Markey MK. Exploring the usability of mobile apps supporting radiologists' training in diagnostic decision making. J Am Coll Radiol 2016 Mar;13(3):335-343. [doi: 10.1016/j.jacr.2015.07.021] [Medline: 26482814]
38. Nielsen J. How to conduct a heuristic evaluation. NN/g Nielsen Norman Group. 1994 Nov 01. URL: https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/ [accessed 2021-11-21]
39. Nielsen J. Why you only need to test with 5 users. NN/g Nielsen Norman Group. 2000 Mar 18. URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2021-11-21]
40. Clarke MA, Schuetzler RM, Windle JR, Pachunka E, Fruhling A. Usability and cognitive load in the design of a personal health record. Health Policy Technol 2020 Jun;9(2):218-224. [doi: 10.1016/j.hlpt.2019.10.002]
41. Flood D, Germanakos P, Harrison R, McCaffery F, Samaras G. Estimating cognitive overload in mobile apps for decision support within the medical domain. In: Marciaszek L, Cuzzocrea A, Cordeiro J, editors. Proceedings of the 14th International Conference on Enterprise Information Systems - Volume 3. Setúbal, Portugal: SciTePress; Jul 28, 2012:103-107.
42. Murphy E, Marquis K, Saner L, Tedesco H, Harris C. Improving electronic data collection and dissemination through usability testing. 1999 Presented at: Federal Committee on Statistical Methodology Research Conference, 1999; November 15-17, 1999; Arlington, Virginia URL: http://www.websm.org/uploadi/editor/Murphy_1999_Improving_electronic.pdf
43. Brown SH, Hardenbrook S, Herrick L, St Onge J, Bailey K, Elkin PL. Usability evaluation of the progress note construction set. Proc AMIA Symp 2001:76-80 [FREE Full text] [Medline: 11825158]
44. Fox J, Mockovak W, Fisher S, Rho C. Usability issues associated with converting establishment surveys to web-based data collection. Federal Committee on Statistical Methodology. 2004. URL: http://www.websm.org/uploadi/editor/1132307602Fox_final.pdf [accessed 2021-11-21]
45. Lai C, Wu C. Display and device size effects on the usability of mini-notebooks (netbooks)/ultraportables as small form-factor Mobile PCs. Appl Ergon 2014 Jul;45(4):1106-1115. [doi: 10.1016/j.apergo.2014.01.009] [Medline: 24518651]
46. Hirschorn DS, Choudhri AF, Shih G, Kim W. Use of mobile devices for medical imaging. J Am Coll Radiol 2014 Dec;11(12 Pt B):1277-1285. [doi: 10.1016/j.jacr.2014.09.015] [Medline: 25467905]
47. Rust RT, Thompson DV, Hamilton RW. Defeating feature fatigue. Harv Bus Rev 2006 Feb;84(2):98-107, 165. [Medline: 16485808]
48. Mazur LM, Mosaly PR, Moore C, Marks L. Association of the usability of electronic health records with cognitive workload and performance levels among physicians. JAMA Netw Open 2019 Apr 05;2(4):e191709 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.1709] [Medline: 30951160]
49. Eschenroeder H, Manzione LC, Adler-Milstein J, Bice C, Cash R, Duda C, et al. Associations of physician burnout with organizational electronic health record support and after-hours charting. J Am Med Inform Assoc 2021 Apr 23;28(5):960-966 [FREE Full text] [doi: 10.1093/jamia/ocab053] [Medline: 33880534]
50. Ehrler F, Haller G, Sarrey E, Walesa M, Wipfli R, Lovis C. Assessing the usability of six data entry mobile interfaces for caregivers: A randomized trial. JMIR Hum Factors 2015 Dec 15;2(2):e15 [FREE Full text] [doi: 10.2196/humanfactors.4093] [Medline: 27025648]
51. Shelby-James TM, Abernethy AP, McAlindon A, Currow DC. Handheld computers for data entry: high tech has its problems too. Trials 2007 Feb 20;8:5 [FREE Full text] [doi: 10.1186/1745-6215-8-5] [Medline: 17309807]

52. Cocosila M, Archer N. Physician interaction with mobile and wireless handheld devices: data entry and retrieval challenges. IJHTM 2006;7(3/4):233. [doi: 10.1504/ijhtm.2006.008434]
53. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
54. Belden JL, Grayson R, Barnes J. Defining and testing EMR usability: Principles and proposed methods of EMR usability evaluation and rating. Healthcare Information and Management Systems Society. 2009 Jun. URL: http://hdl.handle.net/10355/3719 [accessed 2021-11-21]
55. Lowry SZ, Ramaiah M, Patterson ES, Prettyman SS, Simmons D, Brick D, et al. (NISTIR 7804-1) Technical evaluation, testing, and validation of the usability of electronic health records: Empirically based use cases for validating safety-enhanced usability and guidelines for standardization. National Institute of Standards and Technology. 2015 Oct 07. URL: https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.7804-1.pdf [accessed 2021-11-21]
56. Middleton B, Bloomrosen M, Dente MA, Hashmat B, Koppel R, Overhage JM, American Medical Informatics Association. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013 Jun;20(e1):e2-e8 [FREE Full text] [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]
57. Rizvi RF, Marquard JL, Seywerd MA, Adam TJ, Elison JT, Hultman GM, et al. Usability Evaluation of an EHR's Clinical Notes Interface from the Perspective of Attending and Resident Physicians: An Exploratory Study. Stud Health Technol Inform 2017;245:1128-1132. [Medline: 29295278]

Abbreviations

EMR: electronic medical record
PACMAD: People At the Centre of Mobile Application Development
TURF: task, user, representation, function

Edited by A Kushniruk; submitted 10.06.21; peer-reviewed by P Weichbroth, G Ghinea; comments to author 10.07.21; revised version received 02.09.21; accepted 06.09.21; published 20.12.21.

Please cite as:
Choemprayong S, Charoenlap C, Piromsopa K
Exploring Usability Issues of a Smartphone-Based Physician-to-Physician Teleconsultation App in an Orthopedic Clinic: Mixed Methods Study
JMIR Hum Factors 2021;8(4):e31130
URL: https://humanfactors.jmir.org/2021/4/e31130
doi:10.2196/31130
PMID:34931991

Original Paper

An Embodied Conversational Agent in an eHealth Self-management Intervention for Chronic Obstructive Pulmonary Disease and Chronic Heart Failure: Exploratory Study in a Real-life Setting

Silke ter Stal[1,2], PhD; Joanne Sloots[3], MSc; Aniel Ramlal[4], MSc; Harm op den Akker[1,2], PhD; Anke Lenferink[3,4], PhD; Monique Tabak[1,2], PhD

[1]eHealth group, Roessingh Research and Development, Enschede, Netherlands

[2]Biomedical Systems and Signals Group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, Netherlands

[3]Department of Pulmonary Medicine, Medisch Spectrum Twente, Enschede, Netherlands

[4]Department of Health Technology and Services Research, Faculty of Behavioural, Management and Social Sciences, Technical Medical Centre, University of Twente, Enschede, Netherlands

Corresponding Author:
Silke ter Stal, PhD
eHealth group
Roessingh Research and Development
Roessinghsbleekweg 33b
Enschede, 7522AH
Netherlands
Phone: 31 88 0875 777
Email: silke.terstal@utwente.nl

Abstract

Background: Embodied conversational agents (ECAs) have the potential to stimulate actual use of eHealth apps. An ECA's design influences the user's perception during short interactions, but daily life evaluations of ECAs in health care are scarce.

Objective: This is an exploratory, long-term study on the design of ECAs for eHealth. The study investigates how patients perceive the design of the ECA over time with regard to the ECA's characteristics (friendliness, trustworthiness, involvement, expertise, and authority), small talk interaction, and likeliness of following the agent's advice.

Methods: We developed an ECA within an eHealth self-management intervention for patients with both chronic obstructive pulmonary disease (COPD) and chronic heart failure (CHF), which we offered for 4 months. Patients rated 5 agent characteristics and likeliness of following the agent's advice before use and after 3 and 9 weeks of use. The amount of patients' small talk interaction was assessed by log data. Lastly, individual semistructured interviews were used to triangulate results.

Results: Eleven patients (7 male and 4 female) with COPD and CHF participated (median age 70 years). Patients' perceptions of the agent characteristics did not change over time (P>.05 for all characteristics) and only 1 participant finished all small talk dialogues. After 3 weeks of use, the patients were less likely to follow the agent's advice (P=.01). The agent's messages were perceived as nonpersonalized and the feedback as inappropriate, affecting the agent's perceived reliability.

Conclusions: This exploratory study provides first insights into ECA design for eHealth. The first impression of an ECA's design seems to remain during long-term use. To investigate future added value of ECAs in eHealth, perceived reliability should be improved by managing users' expectations of the ECA's capabilities and creating ECA designs fitting individual needs.

Trial Registration: Netherlands Trial Register NL6480; https://www.trialregister.nl/trial/6480

(JMIR Hum Factors 2021;8(4):e24110) doi:10.2196/24110

KEYWORDS

embodied conversational agent; eHealth; self-management; design; daily life evaluation

Introduction

The number of people having a chronic disease, such as diabetes, cancer or chronic obstructive pulmonary disease (COPD), is increasing [1]. COPD is a chronic lung disease that is progressive, and often accompanied by comorbidities, such as chronic heart failure (CHF), that further increase the risk of COPD exacerbations, hospitalizations, mortality, and costs [2,3]. Research shows that paper versions of exacerbation action plans tailored for COPD and comorbidities, embedded in a multifaceted self-management intervention, reduce the duration of COPD exacerbations and the risk of respiratory-related hospitalizations [4].

To further facilitate this chronic disease self-management in daily life, eHealth apps can be used. eHealth apps can provide patients insight into their behavior and disease by symptom monitoring, and patient-tailored and accessible support in their home setting, supervised by their health care professional at a distance [5]. Although such apps seem promising, many eHealth apps face the problem of their actual use decreasing after several weeks by a lack of user engagement [6-8]. Research indicates that a patient's use of eHealth apps is influenced by extrinsic motivation cues, such as stimulation by care professionals and fellow patients [5,9,10]. The majority of existing eHealth apps provide such support in the form of plain text or via a text-based question–answer module, whereas face-to-face interaction remains one of the best ways to communicate health information [11,12].

A different way of providing (motivational) support includes the use of embodied conversational agents (ECAs). ECAs are defined as more or less autonomous and intelligent software entities with an embodiment used to communicate with the user [13]. By face-to-face interaction with the user, ECAs can build trust and rapport, that is, agreement or sympathy between people or groups [14]. By building trust and rapport, they could create a companionship with the user, leading to long-term and continuous use [15] and, thereby, stimulate the actual use of the underlying eHealth app. Just as a human's appearance affects how we evaluate a human, an ECA's appearance affects how we evaluate an ECA. When we interact with another human, or ECA, for the first time, we immediately form initial ideas about the other [16,17]. Furthermore, when we have a positive impression about another human, we tend to interact more with that human. This likely applies to human–agent interaction as well, such that we interact more with ECAs of which we have a positive first impression [16,17].

Thus, ECAs have the potential to promote engagement with eHealth apps. However, a recent review on the design of ECAs for eHealth [18] shows no clear consensus on the design of ECAs for eHealth. More specifically, the review states that emotion and empathic behavior seem to positively affect the user's perception of the agent's characteristics, but that these design features do not necessarily lead to users' behavior change. The review also shows that studies mainly focus on the effect of the ECA design at first glance or after short interaction. But, to gain insight into the possible added value of ECAs in eHealth, it is important to evaluate how the ECAs should be designed for the intended context of long-term use in daily life. Only one study reports on the design of an ECA for eHealth in such a long-term, daily life setting [19]. In this study, a virtual hospital discharge nurse discussed the patient's diagnosis and postdischarge self-care with the patient once a day at his or her hospital bed. In addition, the agent instructed the patient about medication, follow-up appointments, and self-care procedures just before hospital discharge. Questionnaires filled out after the hospital discharge showed that the patient's perceived similarity to this agent was significantly associated with the patients liking the agent and their trust in and desire to continue with the agent. In addition, perceived similarity was associated with the patient's working alliance with the agent—which the authors define as "trust and belief in working with the agent to achieve a therapeutic outcome."

To develop ECAs to support users in self-management of chronic diseases, such as COPD and CHF, more research is necessary on how ECA design affects users' perceptions of an ECA in the intended context of use: a long-term, daily life setting. Research should start in early stages of development of such ECAs, as small-scale eHealth evaluation studies focusing on usability, feasibility, and end-user experience allow researchers to gain detailed information that can be used for further improvement of an eHealth app [20]. The importance of applying user-centered design (UCD; ie, designing with end users instead of for end users by involving them in all stages of the development process) is increasingly being recognized to be valuable in health care [21,22]. By involving users to participate in the early stages of development, technical flaws can be understood and overcome [6] and the technology can be developed in such a way to reach clinical value in follow-up larger-scale studies.

This is a first exploratory study on ECA design for eHealth in a long-term, daily life setting. In this study, an ECA is implemented into an eHealth self-management intervention for patients with COPD and CHF, offered for approximately 4 months. The objective of our study was to investigate how users perceive the design of the ECA over time. In particular, how they perceive the agent's characteristics (friendliness, trustworthiness, involvement, expertise, and authority) and the agent's small talk, and how likely they are to follow the agent's advice.

Methods

Overview

This study was performed as part of the MATCH study. The aim of the MATCH study was to investigate the feasibility of an eHealth self-management intervention for patients with COPD and CHF over a 4-month period. The ECA was implemented into this eHealth self-management intervention. The MATCH study was approved by the Twente Medical Ethical Committee and registered in the Netherlands Trial register (NL6480).

Participants

People were included for participation in the MATCH study if they (1) had a clinical diagnosis of both COPD and CHF; (2)

had at least two COPD or CHF exacerbations or at least one hospitalization for COPD or CHF in the 2 years preceding study entry; (3) were at least one week after prednisolone/antibiotics/furosemide course and hospitalization and at least four weeks after rehabilitation; (4) were at least 40 years of age; (5) were able to understand and read the Dutch language; (6) were able to use a smartphone, tablet, or PC; and (7) provided written informed consent prior to participation. People were excluded from participation if they (1) had terminal cancer or were at the end stage of another serious disease, (2) had another serious lung disease, (3) expected cardiovascular intervention within 3 months, (4) were enrolled in randomized controlled trials or a trial with study medication, (5) were waiting for a heart or lung transplantation, and (6) received renal dialysis.

The eHealth-Supported Self-Management Intervention

The self-management intervention was offered through an app on a tablet (eHealth platform, Roessingh Research and Development, Enschede, the Netherlands) [23] and consisted of the modules listed in Textbox 1, as can be seen in Figure 1.

Textbox 1. Modules of the self-management intervention.

Self-management module

- *Daily symptom diary:* registration of symptoms related to COPD (eg, dyspnea, cough), CHF (eg, weight, edema), and common comorbidities (depression, anxiety) and classification of symptoms in case of symptom deterioration determined by the patient by comparing the symptoms experienced in the last 24 hours with his or her "usual" symptoms on his or her "what are my "usual" symptoms" card. In case of any symptom deterioration, patients were asked to classify each symptom as "normal," "slightly increased," or "significantly increased." The daily symptom diary was connected to a decision-support system that automatically launched self-management advice in case of worsening of the patient's clinical condition (according to symptoms and weight). The automated decision support system was translated from an evidence-based self-management intervention including paper versions of multimorbid exacerbation action plans for patients with COPD and comorbidities [4,24].
- *Action list:* a list of actions containing (1) self-management advice determined by the automated decision-support system (eg, initiate self-treatment, perform relaxation exercises from the *exercise module*, call the case manager). In addition, the list contained (2) reminders to measure weight by a smart scale and (3) reminders to complete questionnaires.
- *Phone numbers:* to contact health care providers for support.
- *Health status:* an overview of a patient's health status during the last week, including an indication of no, slightly increased, or significantly increased symptoms.

Monitoring module

A detailed overview of health status, self-reported symptoms, weight, and received advice.

Inhaler module

Monitoring of and feedback on inhaled medication adherence and technique (add-on sensor for Ellipta Amiko Respiro).

Information module

Presents information about self-management including patients' diseases and healthy behavior [4,24].

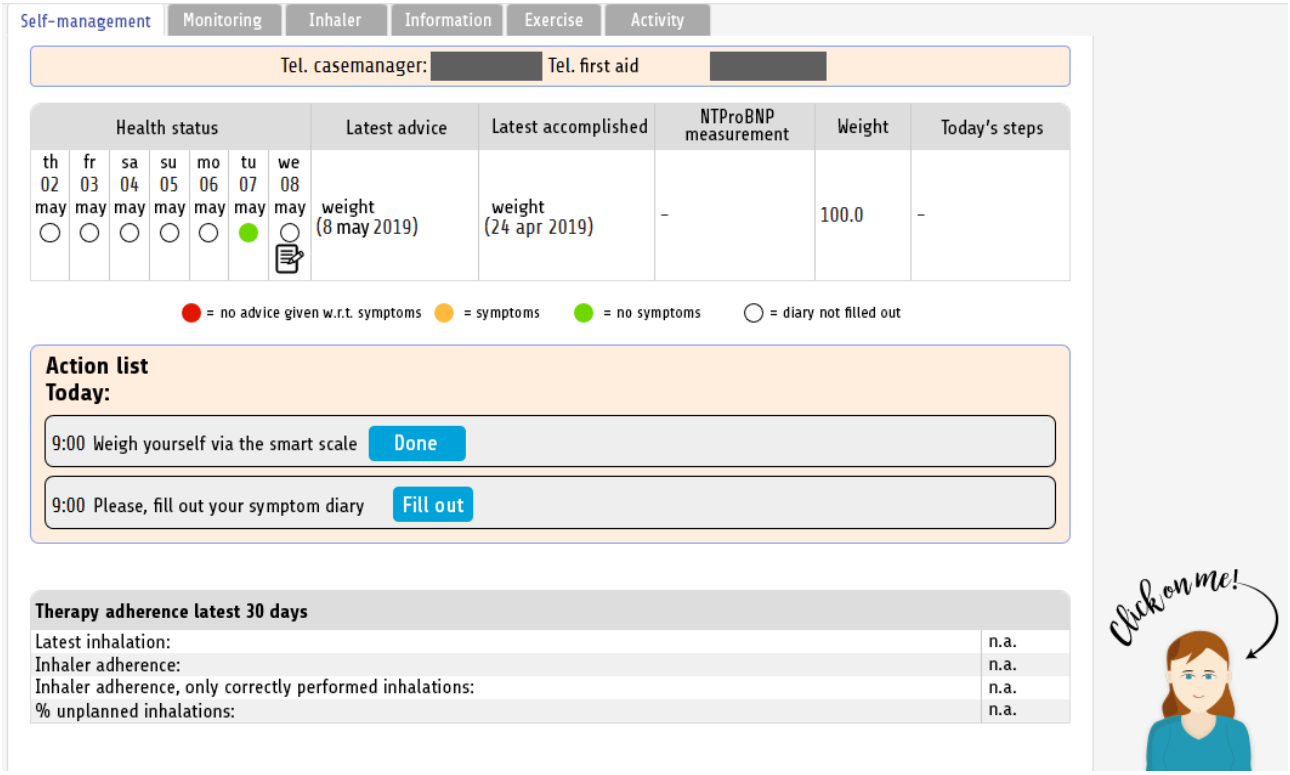
Exercise module

A standardized set of breathing, relaxation, and physical exercises, accompanied by videos and explanation in text.

Activity module

Displays daily physical activity (number of steps measured by the Fitbit Zip).

Figure 1. Home page of the MATCH self-management application, showing the patient's health status and action list, and ECA. The ECA, Sylvia, was always present in the right bottom of the application. The text "click on me" and an arrow pointing to the agent were shown only before the first interaction with the agent started. ECA: Embodied conversational agent



Patients were advised to use the self-management module daily by completing the daily symptom diary, monitoring their weight via the smart scale, and performing the actions on the action list. Furthermore, they were advised to use the inhaler daily. The use of the monitoring, information, exercise, and activity module was voluntarily.

Interaction With Caregivers and Fellow Patients During Self-Management

Patients first attended 3 self-management training sessions (2 group sessions and 1 individual session with the case manager) that among other things included information regarding their diseases and training to recognize symptoms and to practice with using the self-management app. Patients started using the app after the first (group) session, so that questions regarding self-management and the technology could be answered during the next 2 sessions.

For safety during the period of app use, patients were advised to call the case manager (or general practitioner outside office hours) when symptoms did not improve after 2 days of self-treatment and when they experienced dizziness. In addition, the case manager checked health status of the patient (in the app) once per week, and called the patients when they found this was necessary. During the self-management training, patients were instructed that they could call the case manager in case of any questions or doubts. Further, regular health care (eg, visits to their pulmonary physicians and cardiologists) continued as normal during the study.

The Embodied Conversational Agent

The agent characteristics found in the literature were taken into account (Textbox 2; also see Figure 2) when designing the current agent in a creative process with the developers having a description of a persona as outcome.

Textbox 2. Agent characteristics.

Gender: Female

Research indicates that people prefer ECAs that fit their task-conform stereotypes. For health-related tasks (eg, providing medical advice) female agents are preferred [25,26], because these tasks are traditionally being undertaken by women.

Age: Young adult

Research indicates that people prefer young agents over older agents in the context of health, specifically in self-management for chronically ill elderly [27]. As the authors explain, a younger agent might be found more attractive.

Cultural background: Grown up in the Twente region, the Netherlands, living in a terraced house with garden

Research indicates people prefer agents having the same cultural background as themselves [28-30]. The cultural background of the agent is, for example, expressed in the agent's small talk: the agent talks about activities and events related to her place of living.

In addition, to establish a full persona, additional characteristics of the persona were created. Two examples of reasoning behind the characteristics of the persona are given below. The persona used as a guideline to write the dialogues can be seen in Figure 2.

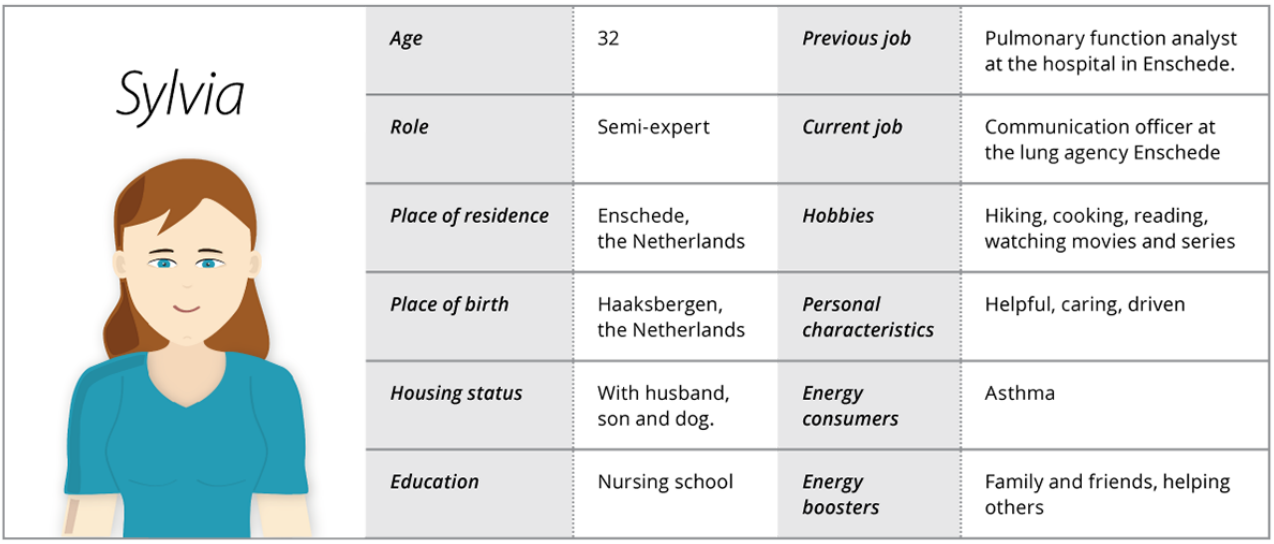
Role: Semiexpert

Because the self-management intervention was supported by a health care professional (nurse practitioner COPD and nurse practitioner CHF), we decided not to create a second medical expert agent. In addition, the goal of the interactive dialogues of the agent was to support patients. Therefore, we gave the agent the role of a "semiexpert," an agent with some experience in chronic diseases (reflected in her career), but that does not act as a doctor or nurse practitioner.

Energy consumers: Asthma

To trigger users to identify with the agent, we decided that the agent has a chronic lung disease as well. However, to ensure the stories of the agent would not become too negative, focusing on limitations related to the disease, we decided the agent has a mild form of asthma.

Figure 2. The persona of the agent Sylvia used as a guideline for writing the dialogues.



The ECA, Sylvia, was implemented into the app via the use of a scalable vector graphics object, including HTML animations. The agent blinked her eyes every 10 seconds, and moved her mouth a fixed period after a new sentence appears on the screen (as if she was talking to the user). The ECA was always present in the right bottom on the pages of the self-management app. Before the user interacted with the agent for the first time, the text "click on me" and an arrow pointing to the agent were shown (Figure 1). After the first interaction, this message disappeared. When hovering over the agent, the agent increased in size.

The content and trigger times of the dialogues were created in collaboration with experts on COPD and CHF to ensure that it was in line with patients' daily practice. Four types of dialogues could be triggered (Textbox 3; also see Figures 3 and 4).

Textbox 3. Types of dialogues.

Action reminders

Dialogues in which the agent reminded the patient of performing actions on the action list of the self-management app (eg, completing the daily symptom diary, weekly questionnaire, or monthly motivation questionnaire; weighing themselves; initiating medication for self-treatment of worsening symptoms; and calling the case manager for support). The agent provided the patient with a general message stating that there were uncompleted actions on the action list, but did not provide the patient with the actual content of these actions.

Inhaler feedback

Dialogues in which the agent informed the patient about (1) the synchronization of the smart inhaler and (2) the inhalation adherence and technique. More specifically, the first type of dialogue informed the user when the smart inhaler had not synced for either 24 or 72 hours. The second type of dialogue informed the user when the inhalation had been skipped for over 2 days, an extra dose had been taken during the last 7 days, the inhalation time of the last inhalation deviated too much from the average duration of the inspiratory flow, and when the position of the device was not optimal.

Health-related tips

Dialogues in which the agent provided the patient with several health-related tips, such as accessing information sources or small actions to perform in daily life. Some of the tips referred to information provided at pages in the self-management app.

Small talk

Chitchat dialogues to increase the patient's engagement [31], stimulating the use of the underlying app. The small talk dialogues were designed as a daily soap series to trigger the patient's curiosity about the continuation of the story. The small talk was split up into 7 "episodes," all containing multiple dialogue steps around a certain theme (the introduction and Sylvia's housing status, husband, child, neighbor, hobbies, and dog). When the patient finished an episode of the small talk, the next episode was unlocked the next day. In the meantime, when the patient clicked on the agent, the agent informed the patient that she does not have time to talk until tomorrow (ie, showed a "wait till tomorrow" message). When the patient finished all 7 episodes, the agent told the patient that she had nothing more to say.

Small talk dialogues could be triggered by the user by clicking on the agent on the home page of the self-management app. The other dialogues were triggered by the system at predefined trigger times:

- *Action reminders:* 1, 2, and 3 hours after an action was added to the action list and not yet performed;
- *Inhaler feedback:* when incorrect inhaler use was measured;
- *Health-related tips:* each day at 15:00 pm;
- *Small talk:* each day at 14:00 pm (only when the patient did not yet initiate a small talk dialogue that day by himself or herself and the small talk was not yet finished).

Each dialogue consisted of one or multiple dialogue steps, containing one or multiple answer possibilities for the user. The agent message was displayed in text and not communicated via speech. An example of the interface of the dialogue step can be seen in Figure 3. Examples of the content of the dialogue steps for every dialogue type can be seen in Figure 4.

Figure 3. Example of a dialogue in which the agent Sylvia reminds the user to perform an action.

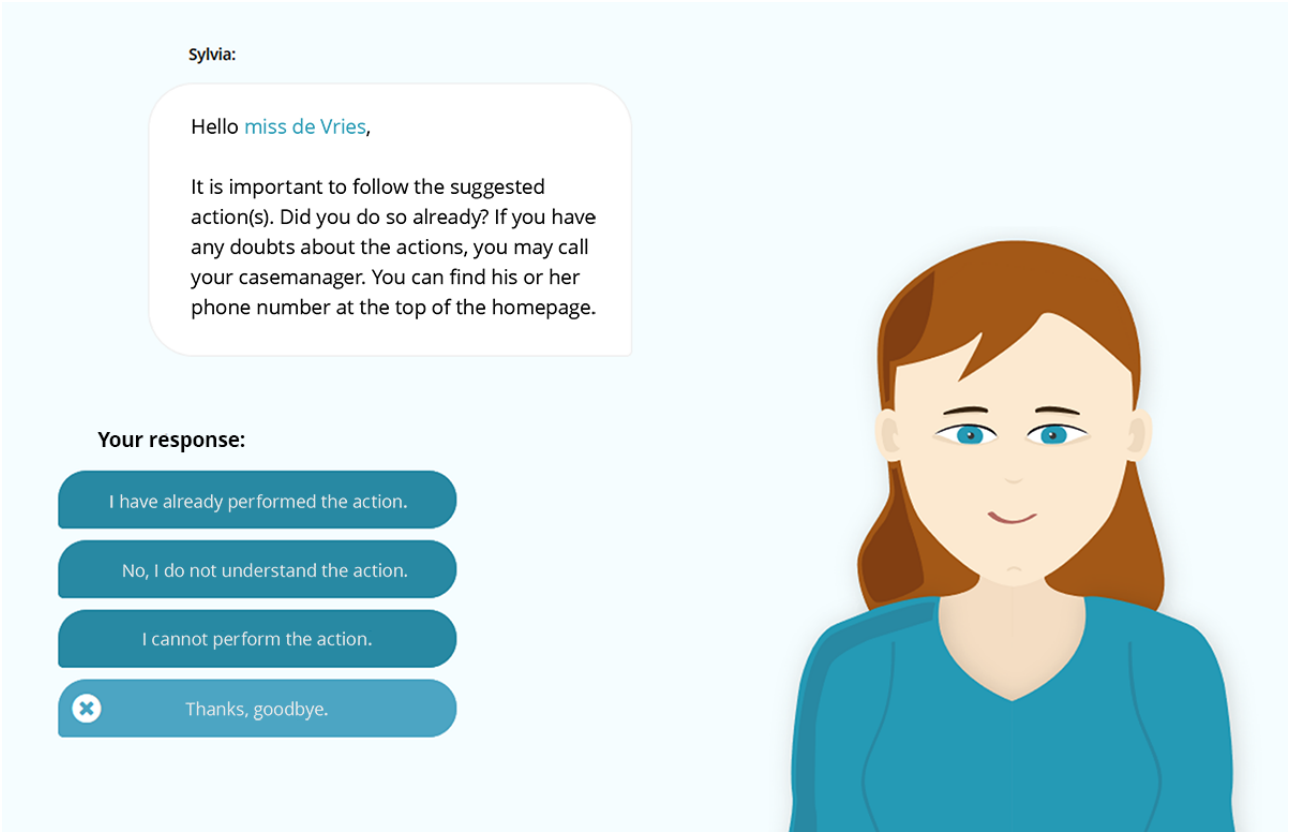


Figure 4. An example dialogue step for every dialogue type that could be triggered.

	Action reminders	Inhaler feedback	Health related tips	Small talk
Message agent 	It is important to follow the suggested action(s). Did you do so already? If you have any doubts about the actions, you may call your casemanager. You can find his or her phone number at the top of the homepage.	You did not take your inhalation medication for two mornings or evenings. It is important not to forget to take your inhalation medication.	On the information page you can find lots of information about various aspects of COPD and heart failure. Have a look if you want to know more about your disease.	*Agent:* Nice to meet you! My name is Sylvia de Ridder. I am 32 years old and I will help you using the application. I will remind you of performing actions, but when you feel like a chat, you are welcome to come by as well!
Reply options patient 	- I have already performed the action. - No, I do not understand the action. - I cannot perform the action. - Thanks, goodbye.	- I will try to improve my habits. - I do not understand how to perform the inhalation. - I do not have any inhalation medication. - Thanks, goodbye.	- Show me the information page. - Thanks, goodbye.	- What did you study? - That sounds great. Are you interested in coaching? - Nice to meet you too. Is your task related to your job? - I agree. Talk to you later.

Procedure

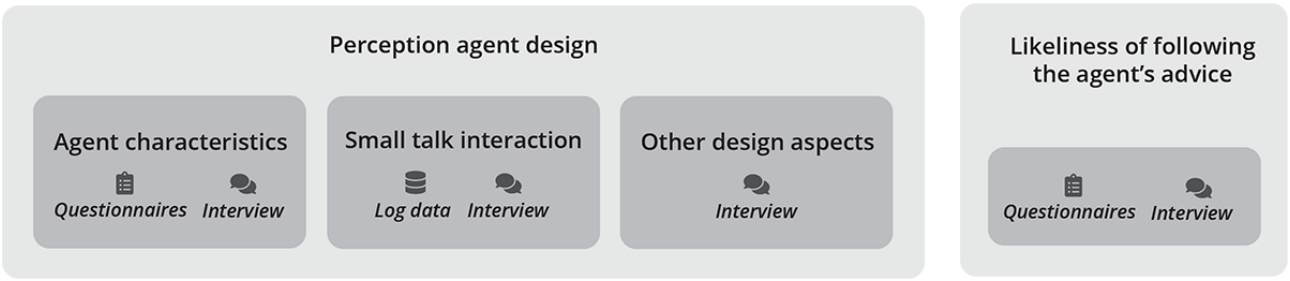
Figure 5 provides an overview of the study procedure. Written informed consent from the participants was obtained prior to study participation. Then, the participants filled out the baseline questionnaire at home (t0). At this point, the participants had not yet seen the app and were not aware of the existence of an agent in the app. The participants were introduced to the agent for the first time in the baseline questionnaire, as a picture of the agent was attached to the questions regarding the agent. The agent was introduced as a hypothetical coach. During the first group session (S1) participants received a tablet, step counter (Fitbit Zip), and smart scale to be used with the app. After this meeting, the participants could already start using the app and sensors were provided. In a second, individual meeting, patients practiced with using the eHealth app according to their individual symptoms (S2). In the second group session (S3), some last questions with respect to self-management and the technology were answered. After the second group session, participants received the add-on sensor for the inhaler and afterward all patients used the app and sensors. Two weeks after the last group session (S3), participants received the intermediate questionnaire (t1). After 9 weeks of use, users received the follow-up questionnaire (t2). Technology usage was logged during the complete period of use. After the end of the use period, the participants were interviewed by an independent interviewer (AR) (t3).

Figure 5. The procedure of the study. S1 = first group session, S2 = individual session, S3 = second group session, t0 = start, t1 = 3 weeks from the start, t2 = 9 weeks from the start, t3 = 15 weeks from the start.

Design and Measurements

We used a mixed-method design, combining both quantitative and qualitative research methods: questionnaires, log data, and semistructured interviews (Figure 6).

Figure 6. Study measurements to evaluate the user's perception of the agent design and the user's likeliness of following the agent's advice.



The patient's perception of the characteristics of the agent and likeliness of following the agent's advice were measured via (1) a baseline questionnaire at t0, (2) an intermediate questionnaire at t1, and (3) a follow-up questionnaire at t2. These paper self-reported questionnaires assessed the patient's perception of:

- Five characteristics of Sylvia (the agent in the MATCH self-management app): friendliness, trustworthiness, involvement, expertise, and authority.
- The importance of these 5 characteristics of an ECA for self-management in general.
- The likeliness of following Sylvia's agent's advice.

The questions on users' perceptions of these characteristics and likeliness of following an ECA's advice were similar to those of 2 other studies [32,33]. All items were assessed by ratings on a 7-point Likert scale. In addition, the baseline questionnaire contained questions related to the patient's characteristics.

Furthermore, small talk interaction was analyzed using (4) log data. The dialogue history of the small talk of the ECA with the patient was logged on the server. For each patient, the date and time of dialogues triggered by either the system or the user were logged. Furthermore, the patient's selected responses were logged per dialogue step of a triggered dialogue, including a date and time.

Lastly, the patient's impression of the agent's characteristics, the likeliness of following the agent's advice, the small talk, and other design aspects were gathered in (5) semistructured interviews.

Data Analyses

Statistical analyses were performed in SPSS 25 (IBM). The respondents' age was treated as a continuous variable, whereas all other respondents' characteristics were treated as categorical variables and responses on Likert scale questions as discrete (ordinal) variables. In the questionnaires, the 5 agent

characteristics were classified as low (a score from 1 until 3), neutral (a score from 3 until 6), or high (a score from 6 until and including 7) on applicability to Sylvia and important characteristic for an ECA for self-management in general. The same classification was used for the user's likeliness of following advice.

For all relations, a related-samples Friedman 2-way analysis of variance by rank was performed as appropriate. The Holm–Bonferroni method was used to correct for multiple comparisons: the comparisons of the ratings for the characteristics of Sylvia and ECA for self-management in general and the likeliness of following Sylvia's advice at t0 (before use), t1 (after 3 weeks of use), and t2 (after 9 weeks of use).

The interviews were transcribed by the interviewer (AR) using automatic transcription in Amberscript and a manual check afterward. Another researcher extracted the interview data focusing on the MATCH agent or ECAs in general (StS). Then, the remaining interview data were thematically analyzed by 2 researchers independently (StS and MT). All themes were grouped either under (1) the patients' perceptions of agent characteristics, (2) small talk interaction, or (3) other design aspects. The themes were coded retrospectively using ATLAS.ti 8, based on the steps proposed in [34]: one researcher (StS) created a first coding scheme and labeled all the data accordingly. A second researcher (MT) used the coding scheme to code a subset of the data. Disagreements between the first and second researcher were discussed and overcome, leading to an updated coding scheme. The first researcher used that updated coding scheme to re-code all data entries and the second researcher then independently re-coded a new subset. Again, disagreements between the 2 researchers were discussed and overcome, leading to the final coding scheme used by the first researcher to re-code all data one final time.

Results

Baseline Demographics

Eleven patients (7 male and 4 female) completed the study procedure until t2, of which 9 agreed to participate in the interview at t3. The age of the participants (n=11) ranged from 49 to 83 years (median 70 years). The highest educational degree for the majority of the participants was high school or vocational education; 1 participant had a university degree. Three participants lived alone, while the others lived with their partner. Four participants indicated that their partner is their informal caregiver, whereas the others said they do not have an informal caregiver. Self-reported tablet skills were high for 4 participants, 3 did not have any experience with a tablet yet, and the rest had some experience.

Patients' Perceptions of Agent Characteristics

Table 1 shows the patients' perceptions of the characteristics of Sylvia and of the important characteristics for an ECA for self-management in general over time. At t0, t1, and t2 Sylvia was rated high on friendliness; on t1 Sylvia was rated high on reliability and low on authority. For all other characteristics, the median rating of the agent was neutral at t0, t1, and t2. In addition, at each point in time, the agent characteristic *authority* was rated neutral on important characteristic for an ECA for self-management. Expertise, reliability, and involvement were rated high on important characteristic for an ECA for self-management in general. Friendliness was rated high on importance at t0 and t2, and neutral on importance at t1.

Table 1. Comparison of the patients' ratings of Sylvia's characteristics and the patients' ratings of the important characteristics for an ECA for self-management in general at t0 (before use), t1 (after 3 weeks of use), and t2 (after 9 weeks of use) using a Friedman 2-way analysis of variance by rank.

Ratings	n	t0, median (IQR)	t1, median (IQR)	t2, median (IQR)	*P* value
Sylvia's characteristics					
Friendliness	9	6.0 (4.0-7.0)	6.0 (4.5-6.0)	6.0 (4.0-7.0)	.45
Expertise	8	5.0 (4.0-7.0)	5.0 (4.0-6.0)	5.0 (4.0-7.0)	.47
Reliability	9	4.0 (4.0-7.0)	6.0 (3.5-6.0)	4.0 (3.8-7.0)	.77
Authority	8	4.0 (2.3-4.8)	2.0 (2.0-5.5)	4.0 (2.3-5.5)	.64
Involvement	8	5.5 (4.0-7.0)	5.0 (4.0-6.0)	4.5 (3.3-7.0)	.68
Important characteristics for an ECA for self-management					
Friendliness	9	6.0 (4.0-7.0)	5.0 (4.0-6.0)	6.0 (4.0-7.0)	.43
Expertise	9	7.0 (7.0-7.0)	7.0 (6.0-7.0)	7.0 (6.5-7.0)	.25
Reliability	9	7.0 (6.5-7.0)	7.0 (6.0-7.0)	7.0 (6.5-7.0)	.84
Authority	9	4.0 (2.0-6.0)	4.0 (1.5-5.5)	4.0 (3.0-4.5)	.65
Involvement	9	7.0 (6.0-7.0)	6.0 (6.0-7.0)	6.0 (4.0-7.0)	.78

In the interviews, the patients commented on some of the above measured characteristics. We identified the themes *friendliness, reliability, expertise,* and *authority*. One participant found the agent (Sylvia) *friendly*, whereas another did not indicate whether Sylvia was friendly, but stressed that an agent for self-management support should be friendly. Furthermore,

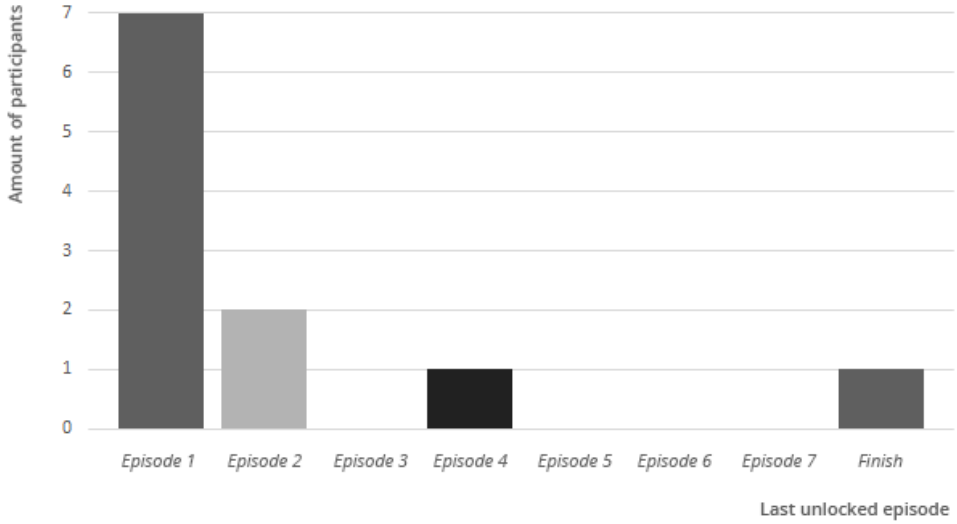
Sylvia was not always seen as *reliable*, supported by a participant indicating that the messages of Sylvia were based on data from a nonreliable Fitbit. Although Sylvia did not provide advice based on the Fitbit data, this participant might have thought this was the case. Another participant indicated that Sylvia sometimes gave tips that did not fit the participant's individual situation; for example, suggesting to perform physical activity when having filled out symptoms in the diary, affecting the agent's reliability. One participant especially commented on the agent's *expertise*, calling Sylvia "a stupid woman." In addition, one participant commented on *authority*, saying

> *She could be your girl next door...If I have medical complaints, I prefer an authority to explain what to do or not to do.*

Small Talk Interaction

Figure 7 shows how many participants unlocked particular episodes, based on the log data. Seven out of the 11 participants did not finish the first episode. Two participants finished the first episode and, therefore, unlocked episode 2. In addition, 1 participant unlocked the episodes until 4. Finally, 1 participant finished all 7 small talk episodes. Two participants were shown a "wait till tomorrow" message, as they already finished a small talk episode that day; one participant saw the message 3 times; and the other 5 times. Finally, the participant that finished all dialogues was shown the message that the small talk was finished for 45 times, meaning this participant clicked on the agent to receive a new small talk dialogue for 45 times, whereas the dialogues were finished.

Figure 7. The number of study participants that unlocked a particular small talk episode.



In the interviews, participants had a few comments on the agent's small talk. One person did not notice that Sylvia talked about her own life. Five participants said they were not interested in the small talk. However, 2 of them thought that people that feel lonely might be interested. One participant (not the participant that finished all the small talk dialogues) showed a more positive attitude toward the agent's small talk:

> *Sylvia could talk nicely, she told me many things, for example, that she was lonely.*

Other Design Aspects

In addition to the agent characteristics and small talk, the analysis of the interviews resulted in the following themes with respect to the agent design: *the agent's appearance*, *frequencies of the messages*, *timing of the messages*, and *the interface design*.

First, *the agent's appearance* was evaluated. One participant preferred to interact with a photo-realistic nurse, instead of a computer-animated figure, because a photo-realistic nurse would make the interaction more personal. The participant also stated:

> *I am not impressed by a cartoon figure.*

Furthermore, the participant described the agent as

> *A male or female such as on the doors of bathrooms.*

Another participant also preferred the agent to look like a nurse; this participant particularly commented on the agent's clothing:

> *Put a white coat and a stethoscope on her.*

The participant described that an agent having a white coat and stethoscope would look more authoritative than the current agent in a t-shirt. Lastly, 1 participant liked that the agent was a woman, because the participant hates listening to men.

Second, 2 participants particularly commented on the *frequency of the messages from the agent*: in their view, they received too many messages. One of them indicated that, therefore, he or she closed the dialogue before reading it.

Third, with respect to the *timing of the messages*, 1 participant would like to receive conformation messages when performing actions (real-time feedback on actions), whereas another suggested that the agent should come back to topics discussed before, as illustrated by:

> *But, then, ask the next day: 'Did you read that? Did you do this?'*

In addition, 2 participants indicated that Sylvia provided unwanted and unsolicited information. One argued that she started to talk about a topic, regardless of whether the user was interested in that topic at a particular moment. This participant, instead, would like to receive the information when asking for

it. The other participant argued that Sylvia provided advice when the participant felt well, whereas this participant only wanted to receive advice when not feeling well. Also, 2 participants said that they did not always have the time to follow the suggested actions or tips when receiving them from the agent. One was really annoyed when receiving messages, like "think about your exercises," straight in the morning, the other explained:

> *When I have to go to work, I do not have time to watch at a 15-minute video.*

Finally, 2 participants found it annoying that the agent already started giving reminders, when opening the app, not having the chance to even perform the action, illustrated by:

> *Look, what really bothered me was that, in the morning, I turned on the device and it [the agent] started with saying: 'Did you follow the instruction?' Well, I did not see any instruction yet.*

Likeliness of Following the Agent's Advice

Table 2 shows the results of the related samples Friedman 2-way analysis of variance by rank, comparing the participant's likeliness of following the agent's advice over time. On t0, t1, and t2 Sylvia scored neutral. A significant difference (P=.01) in the distribution of the values over time was found.

As a second step, pairwise comparisons of t0 and t1, t0 and t2, and t1 and t2 showed no significant difference between t0 and t2 (P=.07) and t1 and t2 (P=.48), but did show a significant difference for the pair t0–t1 (P=.01). The participant's indicated likeliness of following the agent's advice statistically dropped at t1 compared with t0.

Table 2. Comparison of the ratings of participants' likeliness of following Sylvia's advice at t0 (before use), t1 (after 3 weeks of use) and t2 (after 9 weeks of use) using a Friedman 2-way analysis of variance by rank.

Comparison	n	t0, median (IQR)	t1, median (IQR)	t2, median (IQR)	*P* value
Likeliness of following advice	9	6.0 (4.5-7.0)	3.0 (2.0-4.0)	4.0 (2.5-5.0)	.01[a]

[a]Statistical significance is considered if P<.05.

In the interviews, the majority of the participants indicated that they would not *follow the agent's advice*. Two of them questioned the agent's reliability. In line with this, another participant (male) indicated that he would first go to a doctor to verify the agent's advice of taking prednisolone. Although it should be noted that the agent did not provide the patient with advice on taking prednisolone directly, the agent only mentioned that there was an uncompleted action on the action list, which might have been taking prednisolone. However, the actual advice was determined by the automatic decision support system. Another participant argued that the agent did not respond to user input and, therefore, did not find the agent's advice valuable. One participant mentioned not listening to a cartoon figure, and another stated:

> *I do not listen to a device, I do listen to people.*

Furthermore, a participant indicated that adults have their own responsibility, and therefore, this participant did not feel the need for an agent to suggest what to do. One participant argued not having the time to follow the advice, and therefore, not seeing the benefits of the agent's advice. By contrast, 3 participants said they sometimes did follow the agent's advice. One of these participants sometimes performed the physical exercises advised, as this participant valued the exercises. Another indicated to follow the advice of calling the case manager or reading information pages, but would not follow an advice to start prednisolone. The participant said that being wrongly advised to take prednisolone could have negative health consequences, believing that the technology's advice is not always correct. Furthermore, a participant (female) indicated to call the case manager if advised, as she would normally also have done so. As explained before, it should be noted that the actions of calling the case manager and taking prednisolone were part of the action list of the self-management app, but were not presented by the agent itself.

General Attitude Toward the ECA Design

The last theme we identified was *general attitude toward the ECA design*. The theme does not correspond to our main objective, but we present the findings to provide insight into the context of the results described above. The interviews show that the majority of the participants (n=7) did not think that Sylvia had any value, illustrated by comments, such as:

> *I do not have any connection with Sylvia.*
>
> *Sylvia is not it.*

Arguments supporting this opinion were the agent's statements being too obvious, general, or simplistic: a participant described that it was clear that the dialogues were not personalized, but a result of a general set of if–else statements.

Also, Sylvia led to lots of frustration and annoyance, as supported by statements such as:

> *I found this female extremely annoying.*
>
> *Sylvia was a very irritating woman.*

Frustrations were caused by Sylvia providing incorrect feedback on the inhalations and suggesting actions not fitting the user's health status, as illustrated by a participant:

> *I thought: "Gosh, what are you talking about? I'm not complaining about respiratory problems."*

One participant (male) particularly indicated he would like to switch off the agent. By contrast, the interviews showed some positive attitudes toward Sylvia. One participant said that the agent triggered laughing, as Sylvia would adapt the conversation to the answers given. This participant explained:

> *Occasionally, if I felt bad, I could laugh again.*

This participant also said that Sylvia made the app more personal, for example, by addressing the user by his or her first name:

> *It [Sylvia] creates a slightly more informal atmosphere, which I always like, I feel a bit more free.*

In addition, this participant believes people should get used to interacting with agents:

> *When you are at the station, you have this as well [...]. You enter the station and then you face a digital agent. This is something we should get used to, I think.*

Lastly, participants suggested improvements for the interaction with the agent. First, 2 participants explained they would like to be able to type a question in an input field and receive a personalized answer. One of them sketched a scenario in which a patient, who is not feeling well, types in a question into an input field, for example "I am feeling stuffy, but have taken prednisolone: what should I do?" and the app would respond with an answer 24/7. It should be noted that one participant did not understand that Sylvia was a digital agent. He thought that Sylvia represented one of the real people involved in the self-management meetings.

Discussion

Principal Findings

This exploratory study aimed to investigate how an ECA's design is perceived by its users when implemented in a long-term, daily life setting. Although the results of this study should be interpreted carefully, as this is a small-scale study, they provide first insights into an ECA's design for self-management and guidelines for follow-up work in terms of both development and evaluation. Our study shows that the patient's perception of friendliness, expertise, reliability, involvement, and authority of the ECA did not change over time. The majority of the users were not interested in the agent's small talk and the likeliness of following advice decreased after 3 weeks of use.

First, our study shows that the perception of the agent's characteristics at first glance was similar to that after 2 weeks and 9 weeks of use, suggesting that the user's first impression does not change over time. To the best of our knowledge, there are no studies on how these perceptions change over time. But, ter Stal et al [32] showed that an agent's design affects the user's perception after short-term interaction, while Zhou et al [19] showed that this also applies to long-term interaction.

How do you design an agent for self-management that creates positive impressions that persist? Our results suggest that an agent for self-management should be friendly, reliable, and involved and should have expertise, because patients rated these characteristics as important. Cafaro et al [35] found that an agent's friendliness was related to the user's number of agent approaches and likeliness of future encounters with the agent. In addition, the characteristics expertise, reliability, and involvement are found to be important aspects of persuasive systems [36], and eHealth apps in particular [37-39]. However, taking this together does not provide much evidence on what agent characteristics are especially important. In our study, patients gave higher scores for Sylvia's reliability and involvement than for Sylvia's authority. However, patients also indicated that an agent's authority is less important than expertise and reliability. This emphasizes the importance for future ECA design studies to ask for both the perception of the characteristics of the agent designed (ie, the scoring) and the perceived importance of these characteristics for an agent in the specific context. With respect to the agent's authority, our study was indecisive. Different from quantitative data, qualitative data indicated that patients do prefer an agent portraying authority. These contradicting results might be caused by the patients actually meaning that the agent should have expertise, as they indicate in the interviews that "the agent should have authority regarding the topic." Nevertheless, research confirms that people tend to prefer agents designed to fit their task [25,26]. In the context of a self-management intervention for COPD and CHF, we could increase the agent's expertise by having the agent wear a doctor's coat. Whether this actually results in a better perception of the agent should be further investigated.

In addition, our study showed that a photorealistic agent could result in users being more likely to follow the agent's advice, compared with a static cartoon. Van Wissen et al [27] indicated that a more realistic agent appearance increases users' likeliness of following the agent's advice and leads to increased learning of students supported by a pedagogical agent [40]. This increased learning might possibly also apply to a patient's learning about chronic disease self-management. In addition, a realistic agent appearance leads to higher user engagement [26,27,41] and a positive perception of the agent's characteristics, such as its trustworthiness and competence [27]. By contrast, we should avoid the agent being too human-like, as then a mismatch between the users' expectations of the agent and the agent's actual capabilities—a so-called negative adaptation gap—could be created, resulting in the users being disappointed [42]. Future work should investigate the sweet spot between facilitating expertise (through more realism) and managing expectations of intelligence (through reduced realism).

Furthermore, our study showed that the majority of the users was not interested in the agent's small talk. Although we expected that the small talk would increase users' engagement through companionship building with the agent [31], this seemed not the case. A possible explanation might be that the amount of small talk might have exceeded the amount of health-related content, and, therefore, distracted the patients from the actual goal of the app: self-management. We expect that it is better to adapt the amount of small talk to the user, for example, by tracking the user's interaction in the small talk dialogues and adapting the amount of small talk in the future accordingly (ie, users that interact in small talk more often receive small talk more frequently). In addition, the content of the small talk could be adapted to the user. Research shows that tailoring health messages toward personal characteristics pays off [43], suggesting that a user's demographics might affect the type of small talk the user is most engaged by. Future work could focus on how small talk can be personalized to fit the users' personal values and interests.

Lastly, our results show that patients were less likely to follow the agent's advice over time. We expect that the participants had a negative adaptation gap, meaning that their expectations of the agent's capabilities did not match the agent's actual

capabilities [42]. After a few weeks of use, the users might have realized that the agent's messages did not always fit their situation, resulting in a decrease of their likeliness of following the agent's advice. In addition, the agent design led to frustrations, mainly caused by nonpersonalized content and inappropriate feedback, affecting the agent's reliability. Such a mismatch of the content of the agent's message with the user's personal situation was also found by ter Stal et al [33] who evaluated ECAs for health assessment of older adults. Personalizing the agent by providing more specific feedback on user input and health-related data (eg, sensor inhaler data) might improve the likeliness of individuals to follow the agent's advice. However, the technology readiness level (TRL) of the ECA fits the exploratory character of the study, as explained in the staged approach of telemedicine evaluation [20]. In the first stages of an telemedicine evaluation (ie, evaluation of feasibility and user experience), exploratory studies are used to investigate and increase quality of technology, while in later stages, clinical value can be researched with more mature technology in large-scale studies [6,20]. As a consequence, participants' expectations of the technology, especially that of the agent, might have exceeded the functionalities and quality of the technology used. Therefore, we emphasize the importance of managing the participants' expectations of the technology used in a study; that is, they should match the actual TRL of the technology. For an agent in particular, it needs to be explained what the user can expect from the agent, which allows one to focus on the objective of the study.

Our results underpin the importance of applying UCD methods throughout the various development phases of eHealth apps [21]. By incorporating end users early in the development of an ECA for self-management, we learned whether our hypotheses about users' perceptions of the ECA design were correct and gained new insights into how to adapt the design of the ECA to the end users in a next design iteration. The importance of such UCD is recognized more frequently in the field of human–computer interaction, reflected by the development of standards, such as the ISO 9241-210 [44-46]. As described by Mithun et al [44], the ISO 9241-210 clarifies UCD principles and describes that a design process should be iterative; the iteration is the review and refinement of design specifications. Czaja et al [47] stress the importance of UCD for products targeting older adults. They indicate that older adults have unique usability constraints compared with younger adults. As they describe, when usability is improved for older adults, it is also improved for younger adults. Therefore, they stress to take into account the context and characteristics of older adults in the design process. We did so, as many of our participants were older adults.

Strengths and Limitations

The strength of our research is that we evaluated the perception of an agent's design at an early development stage with the end users. Furthermore, participants interacted with an agent in a daily life setting and during a longer period of use: a setting which is rare in agent research, mostly consisting of research on short-term interaction with agents in laboratory settings. However, this long-term, daily life setting put quite some load on participants. Because of the exploratory character of the study, a limited number of patients participated, which should be taken into account when interpreting the results. However, the results can provide guidance for follow-up agent development and evaluation. Furthermore, participants used a Fitbit, smart scale, and a smart sensorized inhaler in combination with the self-management app. Many participants complained about the sensors not working properly. This might have affected the participant's perception of the agent, as some of the messages of the agent were based on incorrect sensor information. Lastly, the interviews focused on all elements of the self-management intervention, not specifically on the design of the agent. Not all participants provided information related to the research question of this study, and, therefore, we should be careful with interpreting the results of the interviews.

Conclusion

This exploratory study provided first insights into ECA design for long-term, daily use. An agent's design is important for patients to establish a good first impression of the agent, which remains during long-term usage. Based on our findings we expect that ECAs do have the potential to be used for self-management, but several design aspects should be investigated in order for ECAs to become successful for increasing engagement in eHealth. When designing ECAs for self-management, we recommend designing an agent that is friendly, reliable, involved, and that has expertise, such that designers can implement and evaluate personalized content and small talk with sufficient variation, and find a good balance between small talk and health-related content. Careful consideration should be given to the apparent realism of the agent to find the sweet spot between facilitating expertise (through more realism) and managing expectations of intelligence (through reduced realism). In combination with managing the user's expectations of the agent capabilities, a personalized ECA design fitting individual needs could increase the agent's reliability and, therefore, the user's likeliness of following the agent's advice. This way, the ECA design could be upgraded to a higher TRL for which the added value and clinical benefits can be evaluated in future research.

Acknowledgments

We thank the patients and care professionals for their participation and feedback. Financial support was provided by the Pioneers in Healthcare innovation fund (2016 MATCH project).

Conflicts of Interest

None declared.

References

1. World Health Organization. World Report on Ageing and Health. 2015. URL: https://www.who.int/ageing/publications/world-report-2015/en [accessed 2020-03-03]
2. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease. 2020. URL: https://goldcopd.org/gold-reports [accessed 2020-03-03]
3. Mannino DM, Thorn D, Swensen A, Holguin F. Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD. Eur Respir J 2008 Oct 14;32(4):962-969 [FREE Full text] [doi: 10.1183/09031936.00012408] [Medline: 18579551]
4. Lenferink A, van der Palen J, van der Valk PD, Cafarella P, van Veen A, Quinn S, et al. Exacerbation action plans for patients with COPD and comorbidities: a randomised controlled trial. Eur Respir J 2019 Nov 14;54(5):180-213 [FREE Full text] [doi: 10.1183/13993003.02134-2018] [Medline: 31413163]
5. Tabak M, VanderValk P, Hermens H, Vollenbroek-Hutten M, Brusse-Keizer M. A telehealth program for self-management of COPD exacerbations and promotion of an active lifestyle: a pilot randomized controlled trial. COPD 2014 Sep;9:935-944. [doi: 10.2147/copd.s60179]
6. Nijland N. Grounding eHealth. Enschede, The Netherlands: University of Twente; 2011.
7. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for Patient Engagement: A Systematic Review. Front Psychol 2015;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
8. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
9. Buimer HP, Tabak M, van Velsen L, van der Geest T, Hermens H. Exploring Determinants of Patient Adherence to a Portal-Supported Oncology Rehabilitation Program: Interview and Data Log Analyses. JMIR Rehabil Assist Technol 2017 Dec 14;4(2):e12 [FREE Full text] [doi: 10.2196/rehab.6294] [Medline: 29242173]
10. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152-e124 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
11. Clark H, Brennan S. Grounding in communication. Perspect. Soc. Shar. Cogn 1991;13:222-233. [doi: 10.1037/10096-006]
12. Bickmore TW, Pfeifer LM, Jack BW. Taking the time to care: Empowering low health literacy hospital patients with virtual nurse agents. 2009 Presented at: International Conference on Human Factors in Computing Systems; 2009; Boston, MA p. 1265-1274. [doi: 10.1145/1518701.1518891]
13. Ruttkay Z, Dormann C, Noot H. Embodied conversational agents on a common ground: A framework for design and evaluation. In: From Brows to Trust. Berlin, Germany: Springer; 2004:27-66.
14. Cambridge University Press. Rapport. URL: https://dictionary.cambridge.org/dictionary/english/rapport [accessed 2019-12-16]
15. Vardoulakis LP, Ring L, Barry B, Sidner CL, Bickmore T. Designing relational agents as long term social companions for older adults. 2012 Presented at: International Conference on Intelligent Virtual Agents; 2012; Santa Cruz, CA p. 289-302. [doi: 10.1007/978-3-642-33197-8_30]
16. Kelley HH. The warm-cold variable in first impressions of persons. J Pers 1950 Jun;18(4):431-439. [doi: 10.1111/j.1467-6494.1950.tb01260.x] [Medline: 15428970]
17. Bergmann K, Eyssel F, Kopp S. A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. 2012 Presented at: International Conference on Intelligent Virtual Agents; 2012; Santa Cruz, CA p. 126-138. [doi: 10.1007/978-3-642-33197-8_13]
18. ter Stal S, Kramer LL, Tabak M, op den Akker H, Hermens H. Design Features of Embodied Conversational Agents in eHealth: a Literature Review. International Journal of Human-Computer Studies 2020 Jun;138:102409. [doi: 10.1016/j.ijhcs.2020.102409]
19. Zhou S, Bickmore T, Paasche-Orlow M, Jack B. Agent-user concordance and satisfaction with a virtual hospital discharge nurse. 2014 Presented at: International Conference on Intelligent Virtual Agents; 2014; Boston, MA. [doi: 10.1007/978-3-319-09767-1_63]
20. Jansen-Kosterink S, Vollenbroek - Hutten MM, Hermens HJ. A renewed framework for the evaluation of telemedicine. 2016 Presented at: eTELEMED 2016: The Eighth International Conference on eHealth, Telemedicine, and Social Medicine; 2016; Venice, Italy.
21. Moody L. User-centred health design: Reflections on D4Ds experiences and challenges. J. Med. Eng. Technol 2015;39(7):395-403. [doi: 10.4324/9781315549026-17]
22. Horne M, Khan H, Corrigan P. People Powered Health: Health For People, By People and With People. 2013. URL: https://media.nesta.org.uk/documents/health_for_people_by_people_and_with_people.pdf [accessed 2020-12-04]
23. op den Akker H, Tabak M, Marin-Perianu M, Huis in 't Veld R, Jones VM, Tönis TM, et al. Development and Evaluation of a Sensor-based System for Remote Monitoring and Treatment of Chronic Diseases. 2012 Presented at: Proceedings of the 6th International Symposium on eHealth Services and Technologies, EHST 2012; 2012; Geneva, Switzerland. [doi: 10.5220/0004474700190027]

24. Lenferink A, Frith P, van der Valk P, Buckman J, Sladek R, Cafarella P, et al. A self-management approach using self-initiated action plans for symptoms with ongoing nurse support in patients with Chronic Obstructive Pulmonary Disease (COPD) and comorbidities: the COPE-III study protocol. Contemp Clin Trials 2013 Sep;36(1):81-89. [doi: 10.1016/j.cct.2013.06.003] [Medline: 23770110]
25. Forlizzi J, Zimmerman J, Mancuso V, Kwak S. How interface agents affect interaction between humans and computers. 2007 Presented at: Designing Pleasurable Product Interfaces; 2007; Helsinki, Finland p. 209-221. [doi: 10.1145/1314161.1314180]
26. Zimmerman J, Ayoob E, Forlizzi J, McQuaid M. Putting a face on embodied interface agents. 2005 Presented at: Designing Pleasurable Products and Interfaces; 2005; Eindhoven, The Netherlands p. 233-248. [doi: 10.1184/R1/6470366.v1]
27. van Wissen A, Vinkers C, van Halteren A. Developing a virtual coach for chronic patients: a user study on the impact of similarity, familiarity and realism. 2016 Presented at: International Conference on Persuasive Technology; 2016; Salzburg, Austria p. 263-275. [doi: 10.1007/978-3-319-31510-2_23]
28. Alsharbi B, Richards D. Using virtual reality technology to improve reality for young people with chronic health conditions. 2017 Presented at: International Conference on Computer and Automation Engineering; 2017; Sydney, Australia p. 11-15. [doi: 10.1145/3057039.3057080]
29. Yin L, Bickmore T, Cortés DE. The impact of linguistic and cultural congruity on persuasion by conversational agents. 2010 Presented at: International Conference on Intelligent Virtual Agents; 2010; Philadelphia, PA p. 343-349. [doi: 10.1007/978-3-642-15892-6_36]
30. Zhou S, Zhang Z, Bickmore T. Adapting a persuasive conversational agent for the Chinese culture. 2017 Presented at: International Conference on Culture and Computing; 2017; Kyoto, Japan p. 89-96. [doi: 10.1109/culture.and.computing.2017.42]
31. Kang SH, Gratch J. People like virtual counselors that highly-disclose about themselves. Stud Health Technol Inform 2011;167:143-148. [Medline: 21685657]
32. ter Stal S, Tabak M, op den Akker H, Beinema T, Hermens H. Who Do You Prefer? The Effect of Age, Gender and Role on Users' First Impressions of Embodied Conversational Agents in eHealth. International Journal of Human–Computer Interaction 2019 Dec 16;36(9):881-892. [doi: 10.1080/10447318.2019.1699744]
33. Ter Stal S, Broekhuis M, van Velsen L, Hermens H, Tabak M. Embodied Conversational Agent Appearance for Health Assessment of Older Adults: Explorative Study. JMIR Hum Factors 2020 Sep 04;7(3):e19987 [FREE Full text] [doi: 10.2196/19987] [Medline: 32886068]
34. Pope C, Mays N. Qualitative methods in health research. In: Qualitative Research in Health Care. Hoboken, NJ: Wiley; 2006:1-11.
35. Cafaro A, Vilhjálmsson HH, Bickmore T, Heylen D, Johannsdottir KR, Valgarosson GS. First impressions: users' judgments of virtual agents' personality and interpersonal attitude in first encounters. 2012 Presented at: International Conference on Intelligent Virtual Agents; 2012; Santa Cruz, CA p. 67-80. [doi: 10.1007/978-3-642-33197-8_7]
36. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: Key Issues, Process Model, and System Features. CAIS 2009;24(28):485-500. [doi: 10.17705/1CAIS.02428]
37. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
38. Van Velsen L, Wildevuur S, Flierman I, Van Schooten B, Tabak M, Hermens H. Trust in telemedicine portals for rehabilitation care: an exploratory focus group study with patients and healthcare professionals. BMC Med Inform Decis Mak 2016 Jan 27;16(1):11-12 [FREE Full text] [doi: 10.1186/s12911-016-0250-2] [Medline: 26818611]
39. Fogg BJ. Computers as Persuasive Social Actors. In: Persuasive Technology. Burlington, MA: Morgan Kaufmann; 2003:89-120.
40. Baylor A, Kim Y. Pedagogical agent design: the impact of agent realism, gender, ethnicity, and instructional role. 2004 Presented at: International Conference on Intelligent Tutoring Systems; 2004; Maceió, Brazil p. 592-603. [doi: 10.1007/978-3-540-30139-4_56]
41. van Vugt H, Konijn E, Hoorn J, Keur I, Eliëns A. Realism is not all! User engagement with task-related interface characters. Interacting with Computers 2007 Mar;19(2):267-280. [doi: 10.1016/j.intcom.2006.08.005]
42. Komatsu T, Kurosawa R, Yamada S. How Does the Difference Between Users' Expectations and Perceptions About a Robotic Agent Affect Their Behavior? Int J of Soc Robotics 2011 Nov 23;4(2):109-116. [doi: 10.1007/s12369-011-0122-y]
43. Beukema S, van Velsen L, Jansen-Kosterink S, Karreman J. "There Is Something We Need to Tell You…": Communicating Health-Screening Results to Older Adults via the Internet. Telemed J E Health 2017 Sep;23(9):741-746. [doi: 10.1089/tmj.2016.0210] [Medline: 28328387]
44. Mithun A, Yafooz WMS. Extended User Centered Design (UCD) Process in the Aspect of Human Computer Interaction. 2018 Presented at: Int. Conf. Smart Comput. Electron. Enterp. (ICSCEE 2018); 2018; Kuala Lumpur, Malaysia p. 1-6. [doi: 10.1109/icscee.2018.8538388]

45. Earthy J, Jones BS, Bevan N. ISO Standards for User-Centered Design and the Specification of Usability. Usablity in Government Systems. 2012. URL: https://www.researchgate.net/publication/288190365_ISO_Standards_for_User-Centered_Design_and_the_Specification_of_Usability [accessed 2021-09-27]
46. International Organization for Standardization. ISO 9241-210:2019. Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2021-02-07]
47. Czaja SJ, Boot WR, Charness N, Rogers WA. Designing for Older Adults. Principles and Creative Human Factors Approaches (Third Edition). Boca Raton, FL: CRC Press; 2019.

Abbreviations

CHF: chronic heart failure
COPD: chronic obstructive pulmonary disease
ECA: Embodied conversational agent
TRL: technology readiness level
UCD: user-centered design

Edited by A Kushniruk; submitted 04.09.20; peer-reviewed by G Huizing, V Stara; comments to author 18.11.20; revised version received 17.12.20; accepted 29.06.21; published 04.11.21.

Please cite as:
ter Stal S, Sloots J, Ramlal A, op den Akker H, Lenferink A, Tabak M
An Embodied Conversational Agent in an eHealth Self-management Intervention for Chronic Obstructive Pulmonary Disease and Chronic Heart Failure: Exploratory Study in a Real-life Setting
JMIR Hum Factors 2021;8(4):e24110
URL: https://humanfactors.jmir.org/2021/4/e24110
doi:10.2196/24110
PMID:34734824

Original Paper

A Mobile App to Support Self-management of Chronic Kidney Disease: Development Study

Talar W Markossian[1], MPH, PhD; Jason Boyda[2], BSc; Jennifer Taylor[1], BA; Bella Etingen[3], PhD; François Modave[4], PhD; Ron Price[2], BA; Holly J Kramer[1,5], MD, MPH

[1]Department of Public Health Sciences, Parkinson School of Health Sciences and Public Health, Loyola University Chicago, Maywood, IL, United States

[2]Department of Informatics and System Development, Loyola University Chicago, Chicago, IL, United States

[3]Center for Innovation for Complex Chronic Healthcare, Hines Veterans Administration Hospital, Hines, IL, United States

[4]Health Outcomes & Biomedical Informatics, College of Medicine, University of Florida, Gainesville, FL, United States

[5]Department of Medicine, Loyola University Chicago, Maywood, IL, United States

Corresponding Author:
Talar W Markossian, MPH, PhD
Department of Public Health Sciences
Parkinson School of Health Sciences and Public Health
Loyola University Chicago
2160 S First Avenue
Maywood, IL, 60153
United States
Phone: 1 7083279027
Email: tmarkossian@luc.edu

Abstract

Background: Chronic kidney disease (CKD) is a common and costly condition that is usually accompanied by multiple comorbidities including type 2 diabetes, hypertension, and obesity. Proper management of CKD can delay or prevent kidney failure and help mitigate cardiovascular disease risk, which increases as kidney function declines. Smart device apps hold potential to enhance patient self-management of chronic conditions including CKD.

Objective: The objective of this study was to develop a mobile app to facilitate self-management of nondialysis-dependent CKD.

Methods: Our stakeholder team included 4 patients with stage 3-4 nondialysis-dependent CKD; a kidney transplant recipient; a caretaker; CKD care providers (pharmacists, a nurse, primary care physicians, a nephrologist, and a cardiologist); 2 health services and CKD researchers; a researcher in biomedical informatics, nutrition, and obesity; a system developer; and 2 programmers. Focus groups and in-person interviews with the patients and providers were conducted using a focus group and interview guide based on existing literature on CKD self-management and the mobile app quality criteria from the Mobile App Rating Scale. Qualitative analytic methods including the constant comparative method were used to analyze the focus group and interview data.

Results: Patients and providers identified and discussed a list of requirements and preferences regarding the content, features, and technical aspects of the mobile app, which are unique for CKD self-management. Requirements and preferences centered along themes of communication between patients and caregivers, partnership in care, self-care activities, adherence to treatment regimens, and self-care self-efficacy. These identified themes informed the features and content of our mobile app. The mobile app user can enter health data including blood pressure, weight, and blood glucose levels. Symptoms and their severity can also be entered, and users are prompted to contact a physician as indicated by the symptom and its severity. Next, mobile app users can select biweekly goals from a set of predetermined goals with the option to enter customized goals. The user can also keep a list of medications and track medication use. Our app includes feedback mechanisms where in-range values for health data are depicted in green and out-of-range values are depicted in red. We ensured that data entered by patients could be downloaded into a user-friendly report, which could be emailed or uploaded to an electronic health record. The mobile app also includes a mechanism that allows either group or individualized video chat meetings with a provider to facilitate either group support, education, or even virtual clinic visits. The CKD app also includes educational material on CKD and its symptoms.

Conclusions: Patients with CKD and CKD care providers believe that a mobile app can enhance CKD self-management by facilitating patient-provider communication and enabling self-care activities including treatment adherence.

(JMIR Hum Factors 2021;8(4):e29197) doi:10.2196/29197

KEYWORDS

chronic kidney disease; mobile app; self-management; mHealth; mobile apps; digital health; kidney disease; smartphone

Introduction

Chronic kidney disease (CKD) affects 15% of the US population and one out of every three US adults is at risk for developing CKD during their lifetime [1]. Proper management of CKD can delay or prevent kidney failure and help mitigate cardiovascular disease risk, which increases as kidney function declines [2]. However, the complexity of CKD management can be overwhelming for patients owing to multiple dietary restrictions, high pill burden, and inadequate disease education. In addition, CKD is usually accompanied by multiple comorbidities including type 2 diabetes, hypertension, and obesity [1,3], and individualized therapies to manage each comorbidity may not be concordant, which may frustrate the patient. Moreover, patients with CKD who receive discordant treatment recommendations owing to conditions that potentially complicate CKD management—for example, health failure or cancer—or because of having multiple medication prescribers have higher risk of health care usage and mortality [4].

Disease self-management is a recognized intervention for improving health status for individuals with chronic conditions [5,6]. While definitions of self-management are heterogeneous [7], the intervention of self-management shifts responsibilities from providers to patients who guide their care in partnership with health care providers [8,9]. Self-management may consist of two overarching domains of care: health care and everyday life [10,11]. Self-management of health care includes the interdependent dimensions of (1) communication between patients and providers, (2) partnership in care, (3) adherence to treatment regimens, (4) self-care activities, and (5) self-care self-efficacy defined as an individual's confidence in disease self-management [12]. Self-management of everyday life involves achieving and maintaining "normality" in usual roles and functioning within the constraints set by living with a chronic condition and its emotional ramifications [10].

Improving self-management of both earlier and late stages of CKD has been associated with better health outcomes [13-18]. In one study, participation in a self-management support program comprising patient education, telephone-based support, and peer support was associated with lower rates of CKD progression [16].

The current pandemic has shown that self-management support needs to be readily accessible, in the hands of patients, and should not require extensive time away from home; this support could potentially be facilitated with a smartphone. In 2021, up to 85% of Americans owned a smartphone, including 61% of adults aged 65 years and older and 76% of persons with household incomes below US $30,000 [19]. In 2017, the number of health-related mobile apps exceeded 318,000, and the number of consumer wearable devices exceeded 340 [20]. In a 2013 survey of US adults, 69% reported keeping track of at least one health indicator, such as weight, diet, exercise routine, or symptoms, through their smartphone [21]. Information technology tools including mobile apps, web-based portals, and web-based educational or coaching interventions are increasingly being adopted to support disease self-management, and growing evidence links their usage to improved clinical outcomes [22-30]. However, mobile apps for CKD self-management targeted either specific clinical or health promotion domains [30,31], did not include the voices of patients with CKD in the app development process, did not focus on earlier stages of CKD, did not receive high ratings for clinical utility or usability, did not address patient safety concerns, and were developed by non-CKD care providers [32].

The overarching objective of this study was to develop a mobile app to enhance kidney disease self-management for nondialysis-dependent CKD. The novelty of our work is in the process of developing a mobile app that supports a holistic care approach and addresses both clinical care and health behavior promotion. In addition, the app is designed by persons with CKD and their providers, including nurses and pharmacists. In subsequent studies, we aim to assess the usability of the mobile app among patients with CKD and to describe patient and provider experiences, and the impact of the mobile app on improving patient activation and cardiovascular health.

Methods

Development Design

The Agile software development methodology guided the overall mobile app design [33]. We emphasized a co-design approach with continuous engagement of stakeholders at every stage of the development process [34] to deliver a product that is as user-friendly as possible. Our stakeholder team included 4 patients with stage 3-4 nondialysis-dependent CKD and a kidney transplant recipient, a caretaker, 2 primary care physicians, 3 PharmDs, a nephrologist, a cardiologist, a registered nurse from the nephrology clinic, a researcher in biomedical informatics (FM), 2 health services and CKD researchers (HK and TM), a systems developer, and 2 programmers. The providers in our key stakeholders' team were recruited from the Loyola University Medical Center (LUMC) and patients were recruited from the LUMC's nephrology clinic. The study was approved by the Loyola University Chicago's institutional research board and patients and providers provided verbal informed consent prior to participation in the focus groups and interviews. All members of the stakeholder team owned a smartphone device.

Data Collection

Two separate focus groups were facilitated by 2 members of the research team (TM and HK) and a programmer (JB). The first focus group included patients and a caretaker, and the second included the registered nurse and the PharmDs. A member of the research team (TM) conducted semistructured interviews with the additional stakeholders (2 primary care physicians, a nephrologist, and a cardiologist). Our focus group and semistructured interviews elicited a list of requirements and preferences regarding the content, features, and technical aspects of the mobile app. Focus groups and interview conversations were semistructured and completed using a guide that was based on CKD knowledge and self-management literature, as well as the mobile app quality criteria from the Mobile App Rating Scale [35]. One of the programmers (JB) reviewed the interview recordings and worked with the other programmer to build the mobile app (based on partial requirements). The programmer (JB) met with the research team biweekly, and intermediate versions of the app were presented to the research team for evaluation and feedback; this early version of the mobile app was then released to our stakeholders.

After the stakeholders had tested the mobile app, we conducted a second round of focus groups and semistructured interviews with stakeholders to elicit feedback about the second mobile app iteration and to identify needed modifications to optimize its usability and usefulness. Stakeholder recommendations for modifications were incorporated into a subsequent version of the mobile app. Because the COVID-19 pandemic began during the study period, the mobile app was modified to include information about COVID-19 infection and prevention. Focus group meetings and interviews were audio-recorded and transcribed verbatim for analysis. Focus groups lasted approximately 120 minutes, and semistructured interviews lasted approximately 35 minutes. The patients were provided lunch and a US $30 gift card to participate in the focus groups. In Table 1, we present the characteristics of the focus group and in-person interview participants (N=13).

Table 1. Characteristics of the participants of focus groups and in-person interviews (N=13). Patients with chronic kidney disease were aged 55-76 years.

Characteristics	Individuals, n
Patients with chronic kidney disease	5
Gender	
Female	1
Male	4
Race	
White	3
Black	2
Hispanic ethnicity	1
Chronic kidney disease stage	
Stage 3	2
Stage 4	2
Kidney transplant recipient	1
Chronic kidney disease care providers	8
PharmDs from population health	3
Registered nurse from the nephrology clinic	1
Physicians	
Primary care	2
Nephrology	1
Cardiology	1

Data Analysis

We used NVivo (version 12, QSR International) qualitative data analysis software to support the analyses of all the focus group and interview data combined. Established qualitative analytic techniques were used, including the constant comparative method [36]. We deductively developed an initial code list a priori, which reflected categories of interest on the basis of elements of our conceptual model, and domains identified by our research team. Within each category, we then inductively developed additional codes and analyzed the text for themes and patterns. Coding entailed an iterative process where our codebook was revised to account for novel instances in the data. We identified key themes and concepts emerging from the data to generate meaningful categorization of the barriers and facilitators of CKD self-management and CKD app content and characteristics [36].

Results

Results Overview

The following themes were identified from the patient and provider stakeholder focus groups and interviews.

Theme A: Need for a Self-management App Specifically Designed for Patients With CKD

Patients described that CKD self-management was multifaceted, and they spoke about the challenges of having multiple apps on their phone, which track the various aspects of CKD self-management (ie, tracking their weight, diet, heart condition, blood pressure, and exercise). As one participant noted, "And what you need to have, which I have never seen, is some form of program that addresses people like us. Who have more than one disease problem." They also described that they have not found an app they like to use to track their blood pressure. The patients in our focus group unanimously expressed that they thought a CKD mobile app could be helpful for disease self-management.

Theme B: Barriers and Respective Facilitators, When Identified, of CKD Self-management

Example extracts from the transcripts about barriers and respective facilitators of CKD self-management are presented in Table 2.

Code B.1: Healthy Lifestyle Is Challenging

Patients and providers agreed that maintaining a healthy lifestyle is challenging, and challenges are amplified in the setting of diabetes. Most providers acknowledged that poor dietary management may exacerbate CKD progression. Meanwhile, patients described the challenges of maintaining a healthy diet in the context of diabetes.

Code B.2: Early-Stage CKD Is Asymptomatic and Early Education Is Key

Most providers described that a major barrier for CKD self-management is the asymptomatic nature of the disease at early stages. Patients became aware of the seriousness of the condition only when the disease progressed to later stages, when it was too late to undo the damage caused to the kidneys.

Code B.3: Social Determinants of Health and the Importance of Addressing Them

Low income, lack of transportation, language, and lack of health literacy are barriers for maintaining healthy behaviors, and availability of translators and social work referrals for outpatient services can help underresourced patients.

Code B.4: Inability to Retain Information Within the Context of a Complicated Condition and Limited Duration of Provider Encounter

Patients described being overwhelmed with the information that the providers share with them during the encounter. However, providers discussed limiting the topics that they address during an encounter because of time constraints and the complexity and urgency of the patients' conditions.

Code B.5: Need for Easily Accessible Nutrition and CKD Education for Patients and Primary Care Providers

Patients described their lack of knowledge about which foods and medications to avoid and the need for simple, easily accessible information. Providers described that patients did not understand how their bodily functions were related, and simple education about the association between high blood pressure, diabetes, and CKD would be effective. Providers also indicated that patients did not understand the medical indications for taking certain medications. Patients also need education on when to contact providers with regard to health concerns. For example, patients should be taught what blood pressure measurement threshold warrants an urgent visit to the doctor's office or the emergency department. The primary care providers described that easily accessible guidelines for early-stage CKD management would be useful for their practice.

Code B.6: The Sequalae of Care Fragmentation and the Importance of Self-advocacy

Patients expressed frustration that they were receiving conflicting health information from different providers and that their providers did not communicate. Providers described similar situations. Providers also expressed frustration about the lack of coverage for medical nutrition therapy (MNT) for patients with early-stage CKD who do not have diabetes or heart disease, who most benefit from MNT to decelerate disease progression. One provider also describes difficulties in finding dietitians with expertise in nondialysis-dependent CKD. Patients described poor communication between different providers and felt this reduced the quality of the care they received.

Code B.7: CKD Is Stigmatized

Two patients spoke about hiding the CKD condition from their spouse and family to not worry them. One patient discussed the stigma associated with CKD and compared it with the stigma associated with obesity.

Code B.8: Feeling Stressed and Staying Positive

Patients spoke about feeling stressed and low on some days, and they described that stress impacted their blood glucose levels and "everything" else. Providers also spoke about the stressful nature of receiving a CKD diagnosis and the importance of maintaining positivity as a motivation for living with a CKD diagnosis.

Code B.9: Patient and Provider Shared Decision-making

Providers described their preference for a shared decision-making process to decide on a clinical course of action and mutual goal-setting as more desirable alternatives than providing authoritative advice. They also described that patients' goals would be most effective when they are mutually set by the patients and providers.

Table 2. Barriers and respective facilitators of chronic kidney disease self-management.

Participant type	Example quotes
Code B.1: Healthy lifestyle is challenging and is amplified in the context of diabetes	
Provider	• *Honestly, I think the number one problem is weight and diet. That's the most difficult thing that we deal with in primary care.* • *If I see an obese, diabetic, hypertensive patient…I urge the patient to lose weight, and I tell them usually THAT's the central goal in order to improve high blood pressure, diabetes and in turn slow down progression of CKD.*
Patient	• *I'm trying to lose weight and when I don't eat, all of a sudden my sugar drops.*
Code B.2: Early-stage chronic kidney disease is asymptomatic and early education is key	
Provider	• *Patients can keep track of protein intake and their sodium intake, but, the majority of my patients are not going to when it's an early disease that have no symptoms and they are not facing any imminent kidney issues.* • *Because they are not having symptom doesn't mean they are not at risk; so I think that point is really important to educate the patients on.*
Patient	• *I think what really opened my eyes is when I came to that point where my creatinine, my GFR, got to 15 and the doctor said you're going to need to go on dialysis.*
Code B.3: Social determinants of health and the importance of addressing them	
Provider	• *I point patients to website to find information. I find that the National Kidney Foundation website has the most patient friendly information. The problem with the CKD population here at…is that health literacy is very low. And getting patients access to information is not always helpful because they are not able to process the information or retain it.*
Code B.4: Inability to retain information within the context of a complicated condition and limited duration of provider encounter	
Provider	• *They come in with chest pain, I can't talk about their kidneys at the same time. At each visit, you have to focus on what is important and you know try to at least touch on many of the other chronic problems as you can.*
Patient	• *You're hearing from this doctor, you're hearing from that doctor. At the end of the day, you don't remember any of that stuff.*
Code B.5: Need for easily accessible nutrition and chronic kidney disease education for patients and primary care providers	
Provider	• *Some patients are confused about the correlation between their blood pressure and kidney disease.* • *Sometimes it hard to stay on top of every specialist's recent guidelines.* • *I deal with every organ system.*
Patient	• *Probably nobody here knew you can't take Ibuprofen. And we are all probably taking it. Like I'm sore from exercising, I'm just going to pop a couple of Ibuprofen.* • *I was doing everything right and then, what I come to find out is that my phosphorus kept staying high. Because salad dressing has a lot of phosphorus in it.*
Code B.6: The sequalae of care fragmentation and the importance of self-advocacy	
Provider	• *Orthopedic doctor may put them on something that they might not recognize as being a nonsteroidal. It happens all the time.* • *People who are sick enough that they are seeing specialists for other diseases whether its heart failure, or diabetes or kidney disease, if they are that sick that they need a specialist, then they are hooked up with nutritionists.*
Patient	• *One doctor was giving me one thing, while the other gave me a medication that hurt my kidneys more.* • *It's up to you, the individual, to have those doctors communicate with each other.* • *I suggested to him that I want to see a kidney specialist.*
Code B.7: Chronic kidney disease is stigmatized	
Patient	• *I don't see any reason why I have to stand up and say 'I am fat'.*
Code B.8: Feeling stressed and staying positive	
Patient	• *And from a personal level stress very much affect my glucose level.* • *It (stress) affects everything.*

Participant type	Example quotes
Provider	• *Now that you've been diagnosed with kidney disease, it's important to maintain positivity…look at that something motivational.*
Code B.9: Patient and provider shared decision-making	
Provider	• *Probably comes out more as an authoritarian: 'You have to monitor these kinds of things.' Rather than trying to work out a sort of agreement with patients.*

Theme C: Visually Appealing App That Is Easy to Use

Patients and providers spoke about wanting an app that is user-friendly, easy to use, and not complicated. Patients also mentioned their preference to self-navigate the app without asking the help of their children. As described in one interview, "You want to go and be able to do this by yourself… I don't want to load too much on there and it starts to get complicated." One patient mentioned her preference for an app in a "language" she understands, "Get something that is in a language that people can understand, and they can see that it's going to help them if they work with it."

Theme D: Peer Support and Scheduled Educational Group Meetings With a Provider

Patients discussed the benefits of sharing strategies for CKD management, including diets and recipes, and recommendations for doctors with their neighbors, and other patients with CKD. One patient reported, "Just about every day we congregate at my house and we talk about these things. We give each other ideas, and doctors, blah blah, blah. We tell them to go on men's MD, researching so we help each other verbally." In contrast, there was mixed enthusiasm for a peer support chat room embedded in the mobile app. Some patients were in favor of chatting with a peer through the app, especially on days when they were "feeling down," while others were opposed to the idea. Providers were generally in favor of the concept of peer support. Nevertheless, both patients and providers supported the concept of periodic support groups focused on CKD education delivered by a provider.

Mobile App Development

Based on these discussions, we used an iterative process to develop a mobile app for CKD self-management. Textbox 1 shows the features included in the mobile app and the care management theme of the mobile app feature [10]. The developed mobile app has a simple user interface where the main functions are accessed through the app's dashboard that provides access to measure-specific (eg, glucose and medications) panels (Figure 1). Each measure's panel provides access to a series on controls that facilitate data collection on the targeted measure. The mobile app user can enter health data including blood pressure, weight, and blood glucose levels; symptoms and severity, where the user will receive a message indicating COVID-19 symptoms or symptoms requiring physician attention. Next, mobile app users can select biweekly goals from a set of predetermined goals with the option to enter customized goals. The user can also keep a list of medications.

Our app includes feedback mechanisms where in-range values for health data are depicted in green and out-of-range values are depicted in red. Certain predetermined goals are linked to biometric data entered by the user and would indicate green if the user met the goal. We ensured that data entered by patients could be downloaded into a user-friendly report that could be emailed or uploaded to an electronic health record, such as Epic, to facilitate communication with providers. The mobile app user can determine which data will be included in the report and dates included. The mobile app also includes a mechanism to facilitate either group or individualized video chat meetings with a provider to facilitate either group support or education or even virtual clinic visits. The CKD app also included educational material on CKD and its symptoms and educated patients on when to contact a provider. The app utilizes native mobile app controls and design principles. The app is implemented with simple language and a large font. The mobile app is available for iOS devices, and we are developing a version for Android devices. Additional details about the mobile app are presented in Textbox 1.

Textbox 1. App characteristics aligned with the self-management of health care dimensions. Since the dimensions of self-management of health care are interdependent, we represented the features of the app under the most relevant dimension.

Communication between patients and caregivers

- Downloadable report of health indicators to share with providers (patients can customize which health indicators to include in the report, along with the monitoring time frame for each indicator) potentially include the following:
  - Log of blood pressure with dates
  - Log of weight with dates
  - Log of symptoms with dates
  - Log of biweekly goals with dates
  - Log of medications currently taking
- Disclosure that the app is a tool to support self-management and will not substitute medical advice or replace physician consultation
- Prompts to contact their provider when entering symptoms which are considered emergent

Partnership in care

- Education content regarding the following:
  - Chronic kidney disease: how kidneys function, recognizing symptoms associated with disease progression, managing blood pressure, medications and supplements to avoid (including herbals, vitamins, and minerals), blood glucose level and implications, chronic kidney disease–related blood analysis
  - Nutrition and recommended diets for patients with chronic kidney disease
  - The concept of self-management and creating SMART (Specific, Measurable, Achievable, Realistic, and anchored within a Time Frame) goals
  - Anxiety and depression
  - COVID-19 symptoms and resources
  - Symptoms most common among patients with chronic kidney disease

Self-care activities

- Setting biweekly goals from a list of pre-existing goals or custom-made goals
- Tracking and monitoring the following health indicators:
  - Blood pressure
  - Weight
  - Blood glucose
- Tracking physical and mental health symptoms including SARS-CoV-2–related symptoms, with a message of when it is recommended to contact their care provider
- Tracking flu and COVID-19 vaccinations
- Tracking and monitoring physical activity (not yet implemented)

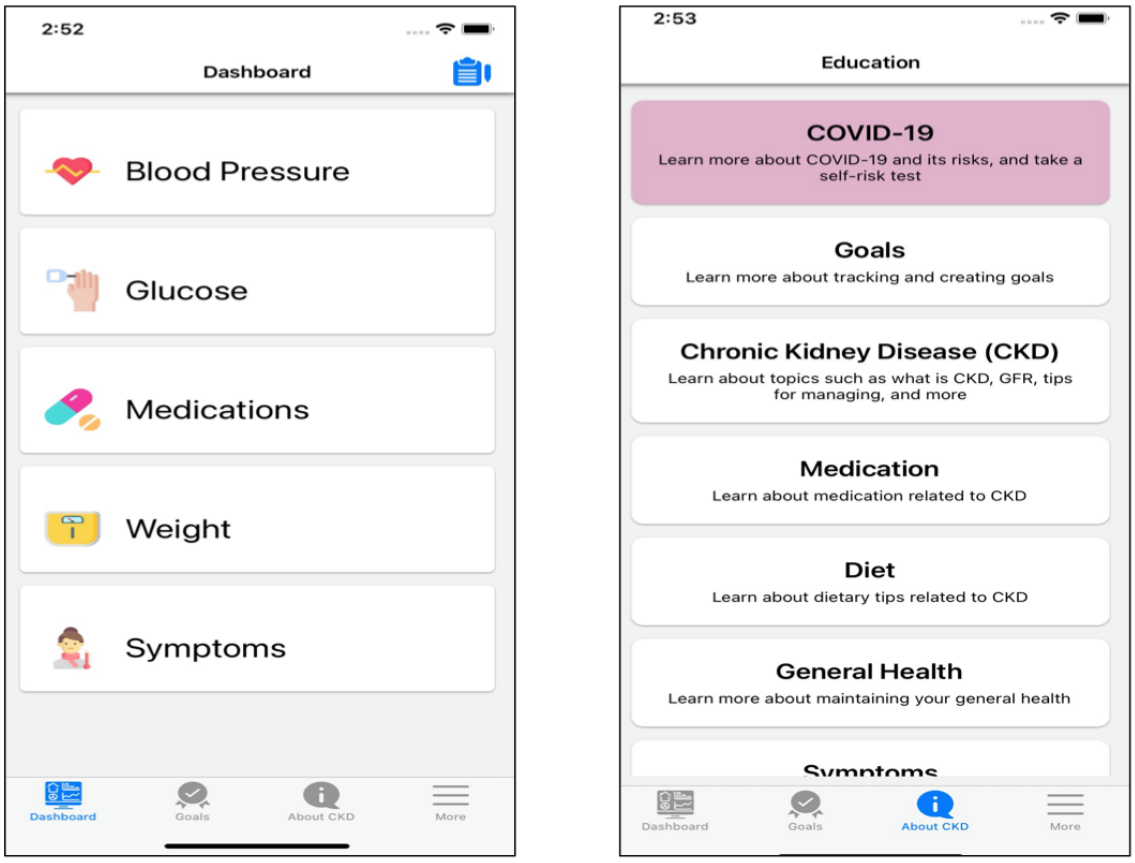
Adherence to medication and treatment regimens

- Medication tracker and alarm; recommend including all medications currently taking prescribed by all providers

Self-care self-efficacy

- Periodic synchronous support group meetings with a focus on education regarding chronic kidney disease facilitated by a provider
- Color-coded reward system (that is linked with health indicators that they tracked) for meeting their biweekly goals (this feature is feasible for the list of pre-existing goals that could be linked to the health indicators)

Figure 1. General and educational dashboards of the chronic kidney disease self-management mobile app.



Discussion

Principal Findings

This study described the iterative process of the development of a mobile app to facilitate self-management of nondialysis-dependent CKD. To our knowledge, our study was the first to describe the process of building a mobile app specific for patients with CKD and is guided by individuals with clinical and methodological expertise. Our stakeholder and production team included patients and providers from the entire spectrum of CKD care, along with health services and bioinformatics researchers. We used an interactive process among the research team, software team, and the stakeholders to elicit preferences and identify the requirements for the mobile app. Our stakeholders identified themes around the barriers and facilitators of CKD self-management, along the themes of CKD self-management of health care domains, which informed the features and content of our app. Our stakeholder patients and providers spoke recurrently about the importance of making the app visually appealing and user-friendly if patients were going to use it. Our findings are overall congruent with a systematic review of information technology solutions used in chronic self-management programs, which revealed that successful solutions included the following key components: education, monitoring, collaboration, and goal-setting [37].

Our findings showed that patients with CKD need a mobile app that is unique for multiple aspects of disease self-management. Unlike other chronic diseases, CKD is usually accompanied by several comorbidities such as diabetes, hypertension, obesity, and heart disease and a mobile app for CKD must address these multiple facets of disease self-management. Several themes emerged from our data that further justifies the need for the CKD app. First, patients need education during the early stages of CKD to prevent disease progression. Second, education and communication with providers need to continue outside of brief clinical encounters to help patients retain information and utilize education on disease self-management. Finally, interventions to improve the health of patients with CKD should include access to MNT. Currently, multiple mobile apps for CKD care management are available and have been previously reviewed [31]. Most existing apps for CKD management were not designed by patients with CKD and their providers, which is consistent for most mobile apps for other chronic diseases [38]. In addition, most mobile apps for CKD do not address both clinical care and health behavior promotion via motivational feedback, goal-setting, or interaction with providers [31].

Our CKD app could be used to enhance both clinical care and health behavior promotion. First, the CKD app provides a mechanism whereby biometric data, medication use, and self-reported symptoms can be tabulated in a report that can be printed or directly uploaded to the electronic health record. Alternatively, patients can simply bring their phone to a clinic visit so that providers may review data on the CKD app with the patient. During a face-to-face or video visit, providers may also work with patients to help them set goals on the app, such as keeping blood pressure at a defined target or losing or maintaining weight over a set period of time. The CKD app may also promote healthy behaviors by providing feedback on biometric data and alerting patients when symptoms may need

urgent discussion with a provider. The CKD app also has a webinar function to facilitate group education or peer support.

In subsequent studies, we will examine the acceptability and usability of the app among patients with early and advanced-stage CKD, and their providers, the impact of using the app on behavioral and clinical outcomes and examine various strategies to integrate the mobile app in the clinical workflow. There is an increasing interest in shared decision-making among patients, payers, and politicians, which was codified by provisions to promote the adoption of decision aids in the 2010 Affordable Care Act [39]. Decision aids help patients become active partners in medical decision-making and include products such as educational booklets, tutorials, and mobile apps.

Strengths and Limitations

The strengths of our study include the use of focus groups and in-depth interviews with patient and provider stakeholders to solicit multiple viewpoints. One of the limitations of our study was the use of small and convenient sample of participants, which may have limited the generalizability of our findings. Subsequent studies among patients with CKD, caregivers, and care providers recruited from a variety of settings would be necessary to assess the acceptability and usability of the mobile app.

Conclusions

Nearly all therapies aimed at preventing kidney disease progression, and decreasing associated complications relies heavily on patient self-management, including recommendations for adherence to medication regimens [40-42], avoidance of further nephrotoxic insults [43] and maintenance of a kidney-friendly diet [44]. A mobile app that integrates domains of clinical care and health behavior promotion may be useful to facilitate CKD self-management, and our mobile app was developed in close collaboration with stakeholders in CKD. Future studies are required to examine the value of the mobile app for CKD self-management.

Acknowledgments

This study was funded by a pilot grant from the Center for Health Outcomes and Informatics Research of Loyola University Chicago. We want to thank the patients and providers for volunteering their time and offering valuable insights. Without their participation, our study would have not been possible.

Conflicts of Interest

None declared.

References

1. Chronic Kidney Disease (CKD) Surveillance System. Centers for Disease Control and Prevention. URL: https://nccd.cdc.gov/CKD/ [accessed 2020-11-18]
2. Making Sense of CKD: A Concise Guide for Managing Chronic Kidney Disease in the Primary Care Setting. National Kidney Disease Education Program. 2014. URL: https://www.niddk.nih.gov/-/media/Files/Health-Information/Health-Professionals/Kidney-Disease/ckd-primary-care-guide-508.pdf [accessed 2021-12-08]
3. Plantinga LC, Boulware LE, Coresh J, Stevens LA, Miller ER, Saran R, et al. Patient awareness of chronic kidney disease: trends and predictors. Arch Intern Med 2008 Nov 10;168(20):2268-2275 [FREE Full text] [doi: 10.1001/archinte.168.20.2268] [Medline: 19001205]
4. Hall RK, Zhou H, Reynolds K, Harrison TN, Bowling CB. A Novel Approach to Developing a Discordance Index for Older Adults With Chronic Kidney Disease. J Gerontol A Biol Sci Med Sci 2020 Feb 14;75(3):522-528 [FREE Full text] [doi: 10.1093/gerona/glz248] [Medline: 31644788]
5. Multiple Chronic Conditions: A Strategic Framework. Optimum Health and Quality of Life for Individuals with Multiple Chronic Conditions. US Department of Health and Human Services. 2010. URL: https://www.hhs.gov/sites/default/files/ash/initiatives/mcc/mcc_framework.pdf [accessed 2021-12-08]
6. Living Well with Chronic Illness: A Call for Public Health Action. Washington, DC: The National Academies Press; 2012.
7. Ryan P, Sawin KJ. The Individual and Family Self-Management Theory: background and perspectives on context, process, and outcomes. Nurs Outlook 2009;57(4):217-225.e6 [FREE Full text] [doi: 10.1016/j.outlook.2008.10.004] [Medline: 19631064]
8. Holman H, Lorig K. Patients as partners in managing chronic disease. Partnership is a prerequisite for effective and efficient health care. BMJ 2000 Feb 26;320(7234):526-527 [FREE Full text] [doi: 10.1136/bmj.320.7234.526] [Medline: 10688539]
9. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [doi: 10.1001/jama.288.19.2469] [Medline: 12435261]
10. Curtin RB, Mapes D, Schatell D, Burrows-Hudson S. Self-management in patients with end stage renal disease: exploring domains and dimensions. Nephrol Nurs J 2005;32(4):389-395. [Medline: 16180780]
11. Clark NM. Management of chronic disease by patients. Annu Rev Public Health 2003;24:289-313. [doi: 10.1146/annurev.publhealth.24.100901.141021] [Medline: 12415147]
12. Bandura A, Freeman WH, Lightsey R. Self-Efficacy: The Exercise of Control. J Cogn Psychother 1999 Jan 01;13(2):158-166. [doi: 10.1891/0889-8391.13.2.158]

13. Donald M, Kahlon BK, Beanlands H, Straus S, Ronksley P, Herrington G, et al. Self-management interventions for adults with chronic kidney disease: a scoping review. BMJ Open 2018 Mar 22;8(3):e019814 [FREE Full text] [doi: 10.1136/bmjopen-2017-019814] [Medline: 29567848]
14. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. Med Care 1999 Jan;37(1):5-14. [doi: 10.1097/00005650-199901000-00003] [Medline: 10413387]
15. Meers C, Singer MA, Toffelmire EB, Hopman W, McMurray M, Morton AR, et al. Self-delivery of hemodialysis care: a therapy in itself. Am J Kidney Dis 1996 Jun;27(6):844-847. [doi: 10.1016/s0272-6386(96)90522-3] [Medline: 8651249]
16. Chen S, Tsai Y, Sun C, Wu I, Lee C, Wu M. The impact of self-management support on the progression of chronic kidney disease--a prospective randomized controlled trial. Nephrol Dial Transplant 2011 Nov;26(11):3560-3566. [doi: 10.1093/ndt/gfr047] [Medline: 21414969]
17. Flesher M, Woo P, Chiu A, Charlebois A, Warburton DER, Leslie B. Self-management and biomedical outcomes of a cooking, and exercise program for patients with chronic kidney disease. J Ren Nutr 2011 Mar;21(2):188-195. [doi: 10.1053/j.jrn.2010.03.009] [Medline: 20650652]
18. Thomas N, Bryar R, Makanjuola D. Development of a self-management package for people with diabetes at risk of chronic kidney disease (CKD). J Ren Care 2008 Sep;34(3):151-158. [doi: 10.1111/j.1755-6686.2008.00032.x] [Medline: 18786082]
19. Mobile Fact Sheet. Pew Research Center. 2021 Apr 07. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/?menuItem=011fca0d-9756-4f48-b352-d58f343696bf [accessed 2021-12-08]
20. The Growing Value of Digital Health. IQVIA. 2017 Nov 07. URL: https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/the-growing-value-of-digital-health.pdf [accessed 2019-04-11]
21. Fox S, Duggan M. Tracking for Health. Pew Research Center. 2013 Jan 28. URL: https://www.pewresearch.org/internet/2013/01/28/tracking-for-health/ [accessed 2021-12-08]
22. Forjuoh SN, Reis MD, Couchman GR, Ory MG. Improving diabetes self-care with a PDA in ambulatory care. Telemed J E Health 2008 Apr;14(3):273-279. [doi: 10.1089/tmj.2007.0053] [Medline: 18570552]
23. van Straten A, Cuijpers P, Smits N. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 2008 Mar 25;10(1):e7 [FREE Full text] [doi: 10.2196/jmir.954] [Medline: 18364344]
24. Lorig KR, Ritter PL, Dost A, Plant K, Laurent DD, McNeil I. The Expert Patients Programme online, a 1-year study of an Internet-based self-management programme for people with long-term conditions. Chronic Illn 2008 Dec;4(4):247-256. [doi: 10.1177/1742395308098886] [Medline: 19091933]
25. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. Med Care 2006 Nov;44(11):964-971. [doi: 10.1097/01.mlr.0000233678.80203.c1] [Medline: 17063127]
26. Song M, Choe M, Kim KS, Yi MS, Lee I, Kim J, et al. An evaluation of Web-based education as an alternative to group lectures for diabetes self-management. Nurs Health Sci 2009 Sep;11(3):277-284. [doi: 10.1111/j.1442-2018.2009.00458.x] [Medline: 19689636]
27. Glasgow RE, Boles SM, McKay H, Feil EG, Barrera M. The D-Net diabetes self-management program: long-term implementation, outcomes, and generalization results. Prev Med 2003 Apr;36(4):410-419. [doi: 10.1016/s0091-7435(02)00056-7] [Medline: 12649049]
28. Grant RW, Wald JS, Schnipper JL, Gandhi TK, Poon EG, Orav EJ, et al. Practice-linked online personal health records for type 2 diabetes mellitus: a randomized controlled trial. Arch Intern Med 2008 Sep 08;168(16):1776-1782 [FREE Full text] [doi: 10.1001/archinte.168.16.1776] [Medline: 18779465]
29. The GVODH. Evidence and Impact on Human Health and the Healthcare System Institute Report, Nov 07, 2017. Evidence and Impact on Human Health and the Healthcare System Institute Report. 2017 Nov 07. URL: https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health [accessed 2020-12-01]
30. Ong SW, Jassal SV, Miller JA, Porter EC, Cafazzo JA, Seto E, et al. Integrating a Smartphone-Based Self-Management System into Usual Care of Advanced CKD. Clin J Am Soc Nephrol 2016 Jun 06;11(6):1054-1062 [FREE Full text] [doi: 10.2215/CJN.10681015] [Medline: 27173169]
31. Siddique AB, Krebs M, Alvarez S, Greenspan I, Patel A, Kinsolving J, et al. Mobile Apps for the Care Management of Chronic Kidney and End-Stage Renal Diseases: Systematic Search in App Stores and Evaluation. JMIR Mhealth Uhealth 2019 Sep 04;7(9):e12604 [FREE Full text] [doi: 10.2196/12604] [Medline: 31486408]
32. Singh K, Diamantidis CJ, Ramani S, Bhavsar NA, Mara P, Warner J, et al. Patients' and Nephrologists' Evaluation of Patient-Facing Smartphone Apps for CKD. Clin J Am Soc Nephrol 2019 Apr 05;14(4):523-529 [FREE Full text] [doi: 10.2215/CJN.10370818] [Medline: 30898873]
33. Martin R. Agile Software Development, Principles, Patterns, and Practices. London: Pearson Education; 2003.
34. Jessen S, Mirkovic J, Ruland CM. Creating Gameful Design in mHealth: A Participatory Co-Design Approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]
35. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]

36. Strauss A, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, CA: Sage Publications; 1998.
37. Solomon MR. Information Technology to Support Self-Management in Chronic Care. Dis Manag Health Out 2008 Mar 21;16(6):391-401. [doi: 10.2165/0115677-200816060-00004] [Medline: 5637250]
38. Adu MD, Malabu UH, Callander EJ, Malau-Aduli AE, Malau-Aduli BS. Considerations for the Development of Mobile Phone Apps to Support Diabetes Self-Management: Systematic Review. JMIR Mhealth Uhealth 2018 Jun 21;6(6):e10115 [FREE Full text] [doi: 10.2196/10115] [Medline: 29929949]
39. Braddock CH. The emerging importance and relevance of shared decision making to clinical practice. Med Decis Making 2010;30(5 Suppl):5S-7S. [doi: 10.1177/0272989X10381344] [Medline: 20881148]
40. Kidney Disease Outcomes Quality Initiative (K/DOQI) Group. K/DOQI clinical practice guidelines for management of dyslipidemias in patients with kidney disease. Am J Kidney Dis 2003 Apr;41(4 Suppl 3):I-IV, S1. [Medline: 12671933]
41. Kidney Disease Outcomes Quality Initiative (K/DOQI). K/DOQI clinical practice guidelines on hypertension and antihypertensive agents in chronic kidney disease. Am J Kidney Dis 2004 May;43(5 Suppl 1):S1-290. [Medline: 15114537]
42. K/DOQI Workgroup. K/DOQI clinical practice guidelines for cardiovascular disease in dialysis patients. Am J Kidney Dis 2005 Apr;45(4 Suppl 3):S1-153. [Medline: 15806502]
43. Harkonen S, Kjellstrand C. Contrast nephropathy. Am J Nephrol 1981;1(2):69-77. [doi: 10.1159/000166493] [Medline: 6758574]
44. Clinical practice guidelines for nutrition in chronic renal failure. K/DOQI, National Kidney Foundation. Am J Kidney Dis 2000 Jun;35(6 Suppl 2):S17-S104. [doi: 10.1053/ajkd.2000.v35.aajkd03517] [Medline: 10895784]

Abbreviations

CKD: Chronic Kidney Disease
LUMC: Loyola University Medical Center
MNT: medical nutrition therapy
SMART: Specific, Measurable, Achievable, Realistic, and anchored within a Time Frame

Edited by A Kushniruk; submitted 29.03.21; peer-reviewed by C Diamantidis, M Singh, N Koizumi; comments to author 11.06.21; revised version received 16.07.21; accepted 06.11.21; published 15.12.21.

Please cite as:
Markossian TW, Boyda J, Taylor J, Etingen B, Modave F, Price R, Kramer HJ
A Mobile App to Support Self-management of Chronic Kidney Disease: Development Study
JMIR Hum Factors 2021;8(4):e29197
URL: https://humanfactors.jmir.org/2021/4/e29197
doi:10.2196/29197
PMID:34914614

Original Paper

Perceptions of Older Men Using a Mobile Health App to Monitor Lower Urinary Tract Symptoms and Tamsulosin Side Effects: Mixed Methods Study

Elizabeth Y Wang[1], MD, MPH; Benjamin N Breyer[2], MD, MAS; Austin W Lee[2], MD; Natalie Rios[2], BS; Akinyemi Oni-Orisan[2], PharmD, PhD; Michael A Steinman[2], MD; Ida Sim[2], MD, PhD; Stacey A Kenfield[2*], ScD; Scott R Bauer[2*], MD, MS

[1]Columbia Vagelos College of Physicians and Surgeons, New York, NY, United States

[2]University of California San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Scott R Bauer, MD, MS
University of California San Francisco
550 16th St, 6th floor
Box 1695
San Francisco, CA, 94121
United States
Phone: 1 4152214810 ext 24322
Email: Scott.Bauer@ucsf.edu

Related Article:

Companion article: https://formative.jmir.org/2021/12/e30762

Abstract

Background: Mobile health (mHealth) apps may provide an efficient way for patients with lower urinary tract symptoms (LUTS) to log and communicate symptoms and medication side effects with their clinicians.

Objective: The aim of this study was to explore the perceptions of older men with LUTS after using an mHealth app to track their symptoms and tamsulosin side effects.

Methods: Structured phone interviews were conducted after a 2-week study piloting the daily use of a mobile app to track the severity of patient-selected LUTS and tamsulosin side effects. Quantitative and qualitative data were considered.

Results: All 19 (100%) pilot study participants completed the poststudy interviews. Most of the men (n=13, 68%) reported that the daily questionnaires were the right length, with 32% (n=6) reporting that the questionnaires were too short. Men with more severe symptoms were less likely to report changes in perception of health or changes in self-management; 47% (n=9) of the men reported improved awareness of symptoms and 5% (n=1) adjusted fluid intake based on the questionnaire. All of the men were willing to share app data with their clinicians. Thematic analysis of qualitative data yielded eight themes: (1) orientation (setting up app, format, symptom selection, and side-effect selection), (2) triggers (routine or habit and symptom timing), (3) daily questionnaire (reporting symptoms, reporting side effects, and tailoring), (4) technology literacy, (5) perceptions (awareness, causation or relevance, data quality, convenience, usefulness, and other apps), (6) self-management, (7) clinician engagement (communication and efficiency), and (8) improvement (reference materials, flexibility, language, management recommendations, and optimize clinician engagement).

Conclusions: We assessed the perceptions of men using an mHealth app to monitor and improve management of LUTS and medication side effects. LUTS management may be further optimized by tailoring the mobile app experience to meet patients' individual needs, such as tracking a greater number of symptoms and integrating the app with clinicians' visits. mHealth apps are likely a scalable modality to monitor symptoms and improve care of older men with LUTS. Further study is required to determine the best ways to tailor the mobile app and to communicate data to clinicians or incorporate data into the electronical medical record meaningfully.

(JMIR Hum Factors 2021;8(4):e30767) doi:10.2196/30767

KEYWORDS

BPH; mobile health; mHealth; telehealth; telemedicine

Introduction

Lower urinary tract symptoms (LUTS) comprise a complex and heterogenous syndrome, including urinary urgency, urinary frequency, weak urinary stream, hesitancy, straining, incomplete bladder emptying, nocturia, and urinary incontinence [1]. LUTS are chronic and progressive [2], and they affect as many as 72% of men over the age of 40 years [3]. First-line treatment for LUTS due to benign prostatic hyperplasia (BPH) includes medical management with alpha-1-andrenergic receptor blockers, such as tamsulosin. Although some men achieve symptomatic relief with this regimen, many do not, and placebo effects make it difficult for patients and their providers to determine true clinical response [4]. More than 10% of men taking tamsulosin will also suffer from undesirable side effects, such as dizziness, orthostatic hypotension, headache, sexual dysfunction, and rhinitis. In addition to contributing to unnecessary polypharmacy, these potential side effects can be debilitating and a major concern for patients with LUTS, especially in older men or men who achieve minimal or no symptomatic relief from tamsulosin [5]. Thus, successful management of LUTS requires a balance of both benefits and harms from interventions, including tamsulosin, which are highly variable between individuals and may not be accurately communicated or recorded during routine medical encounters. Lowering barriers to improve communication of symptoms and medication side effects may improve understanding and adherence in a significant proportion of men who are incorrectly labeled as refractory to medical LUTS management. In addition, mobile apps may increase patients' awareness of variability in symptom severity, identify triggers of symptoms or side effects, assist clinicians by quantifying the efficacy and adverse effects of LUTS interventions, and ultimately facilitate shared decision making based on whether a medication is continuing to provide net benefit to the patient.

Mobile apps are increasingly used to monitor symptoms and side effects in a wide range of urologic conditions. For men with LUTS, mobile apps have been successfully used to administer questionnaires [6], increase peak flow rate (ie, with the sound of running water) [7], and guide clinical decision making [8]. Given these findings, a well-designed and thoughtfully implemented mobile app could streamline the way that patients report symptoms and improve characterization of LUTS severity and variability. In addition, mobile health (mHealth) apps may provide an efficient mechanism to generate a repository of patient data that could be used to identify new LUTS phenotypes, define treatment response or harms, and ultimately help clinicians tailor medical management.

To our knowledge, the use of mobile apps to track symptoms and side effects in men with LUTS has not been studied using qualitative methodology. Qualitative studies are uniquely equipped to explore the patient perspective and help ensure that future intervention designs are grounded in the patient experience. Thus, we designed the Placebo-Controlled, Randomized, Patient-Selected Outcomes, N-of-1 Trials (PERSONAL) pilot study, a 2-week intervention to determine the feasibility and acceptability of daily LUTS severity and tamsulosin side-effect assessment through a mobile app among older men with LUTS receiving chronic tamsulosin therapy. Following this study period, we aimed to explore the men's insights about using an mHealth app to track their symptoms and medication side effects via interviews. Here, we present the findings and implications of the first mixed methods study exploring the experience of men with LUTS after using an mHealth app to track their symptoms and medication side effects.

Methods

Study Design

A convenience sample of 19 men was recruited from an academic urology clinic for the PERSONAL pilot study. Recruitment was targeted at men with LUTS who may be unsure if the benefits of tamsulosin outweighed the harms, specifically older men who were both receiving chronic tamsulosin therapy and interested in tracking their daily urinary symptoms and tamsulosin side effects outside of regular clinic visits. Men were eligible if they were taking tamsulosin for LUTS for at least 12 months, received care from a urologist at our institution, and had previously consented to being contacted for research purposes via the electronic medical record. The sample size was based on feasibility of recruitment. Additional inclusion and exclusion criteria are outlined in Multimedia Appendix 1. First, the participants completed a baseline survey describing their LUTS severity and medication side effects; in the same survey, they selected up to three individual symptoms and three tamsulosin side effects to track daily for 2 weeks. Then, the participants underwent an orientation phone call to set up the PERSONAL mobile app; they were guided on how to use the app and had the opportunity to ask questions. The participants received daily questionnaires through the mobile app at a prespecified time of their choice, and the questionnaire data from all participants were then collected in a secure cloud-based database made accessible to the research team; participants could review their own results throughout the study period. Additional study design details were previously published [9], and results pertaining to the PERSONAL pilot study itself will be published in a separate manuscript.

Interviews

To gain a more nuanced understanding of participants' experiences using the PERSONAL mobile app following the 2-week data collection period, semistructured feasibility interviews were administered by telephone, ranging from 10 to 30 minutes. The interviewer (EYW) was a medical trainee and research associate with formal qualitative research training, who had no interactions with the participants prior to the interviews.

The interview guide is available in Multimedia Appendix 2. Responses were recorded in REDCap (Research Electronic Data Capture) and field notes were made afterward. Due to the relatively structured nature of interview questions, qualitative responses were comprehensively transcribed in REDCap during the phone interviews, but were not audio-recorded. EYW manually coded these open-ended responses using a data-driven approach; codes were iteratively refined into themes. EYW reviewed codes and themes intermittently with coauthors SAK and SRB.

Statistical Approach

Continuous variables were reported using median and IQR, and categorical data were summarized using frequencies. Thematic analysis of open-ended questions yielded 22 codes, which were refined into eight themes. EYW completed initial coding; codes and themes were finalized in discussion with coauthors SRB and SAK. Qualitative data were integrated relative to quantitative responses [10]. The study is reported in concordance with the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist [11].

Results

All 19 participants who participated in the 2-week PERSONAL pilot study completed poststudy interviews; their demographics and self-selected symptoms and side effects are reported in Table 1. All participants found the setup process easy via phone orientation; the majority of participants (n=15, 79%) reported no issues at all, and 4 participants (21%) mentioned minor initial issues, which were easy to resolve (device incompatibility, asked partner to help download app, etc). Some participants suggested an online video or a written document in addition to the phone orientation. Stated reasons for wanting a written document included lower English proficiency and to have materials for future reference.

About two-thirds of participants felt that the questionnaires were just the right length, and about one-third felt that they were too short; the latter participants reported concerns about adequate detail or the appropriate gradations used in the questions and answer choices, commenting that those with milder symptoms might prefer additional gradations on the lower end of the spectrum. Some participants reported frustration about the side-effects questions, as they had difficulty parsing out which side effects to specifically attribute to tamsulosin. A few men were willing to log answers for seven to eight symptoms and side effects rather than the recommended limit of three. About two-thirds of participants set app notifications for mornings, which they found particularly useful for reporting nocturnal symptoms, such as nocturia.

Some participants reported that they did not receive daily notifications via the mobile app due to technical difficulties, but the large majority of participants completed the questionnaire by “habit” at the same time every day—no participants logged responses purely as an immediate response to bothersome symptoms. Out of 19 participants, 9 (47%) felt that this app changed their perception of their health or LUTS management; 8 of these 9 patients (89%) reported increased awareness of symptoms and 1 participant (11%) was able to adjust his fluid intake to improve LUTS. All participants would be willing to share data with their clinician to complement their usual care.

Almost half (8/19, 42%) of the participants reported prior use of other health apps or devices. When prompted to compare those experiences with the study app, most agreed that this app had a simpler interface than others they had encountered, though 1 participant (5%) noted that the app may still be inaccessible in populations with lower English proficiency.

Table 1. Participant demographics and responses to structured interviews.

Characteristic	Value (N=19)
Age (years), median (IQR)	70 (62-75)
Race, n (%)	
White	13 (68)
Black	0 (0)
Asian	3 (16)
Other	3 (16)
Ethnicity, n (%)	
Not Hispanic, Latino, or Spanish	16 (84)
Mexican, Mexican American, or Chicano	1 (5)
Puerto Rican	1 (5)
Another Hispanic, Latino, or Spanish origin	1(5)
Preferred orientation format, n (%)	
Online video	10 (53)
On-demand phone support	6 (32)
In-person demonstration	4 (21)
Phone orientation	3 (16)
Length of daily questionnaire, n (%)	
Took more time than it should	0 (0)
About right	13 (68)
Took less time than it should	6 (32)
Time of day questionnaire usually completed, n (%)	
Morning	13 (68)
Afternoon	1 (5)
Evening	4 (21)
Sporadic	1 (5)
Frequency of app usage without alert or notification, n (%)	
Never	6 (32)
A few times per week	11(58)
More than half of the days per week	1 (5)
Once per day	2 (11)
2-3 times per day	0 (0)
More than 3 times per day	0 (0)
Usefulness of urinary symptom assessment[a], median (IQR)	4 (4-4)
Usefulness of medication side-effect assessment[a], median (IQR)	4 (2-5)
Reported change in urinary symptom management due to the use of the PERSONAL[b] app (during the 2-week study), n (%)	9 (47)
Would allow clinician to see PERSONAL app data, n (%)	19 (100)
Reported use of other health apps or devices, n (%)	8 (42)
Other health apps or devices, n (%)	
Fitbit	4 (21)
Other[c]	6 (32)

[a]On a scale of 1 to 5, where 1 is "not at all useful" and 5 is "very useful."

[b]PERSONAL: Placebo-Controlled, Randomized, Patient-Selected Outcomes, N-of-1 Trials.

[c]Other health apps or devices include Fitbit-like device, Apple Watch, MyFitnessPal, Brainscape, and the iPhone heart rate monitor.

The qualitative analysis generated eight themes (ie, orientation, triggers, daily questionnaire, technology literacy, perceptions, self-management, clinician engagement, and improvement) and 22 codes organized into three categories (ie, mHealth app, clinical use, and next steps) whose relationships are illustrated in Figure 1. Definitions for themes and select quotations organized by code are included in Table 2. For example, under the theme "perceptions" and code "usefulness," one participant said, "I learned some things that I didn't know about my condition before; I was surprised about what I learned about my symptoms—'cause then I can talk to my doctor about it." Under the theme "clinician engagement" and code "efficiency," another participant said, "I think it would help that we would be looking at this issue a little more closely as opposed to a casual conversation during the visit. I think the doctor would be able to monitor the frequency of symptoms I'm having either throughout the night or the day and we can decide if the dosage needs to be changed." Additional quotations are presented in Multimedia Appendix 3.

Figure 1. Codes and themes generated from the qualitative analysis. Orange boxes represent themes and blue boxes represent codes.

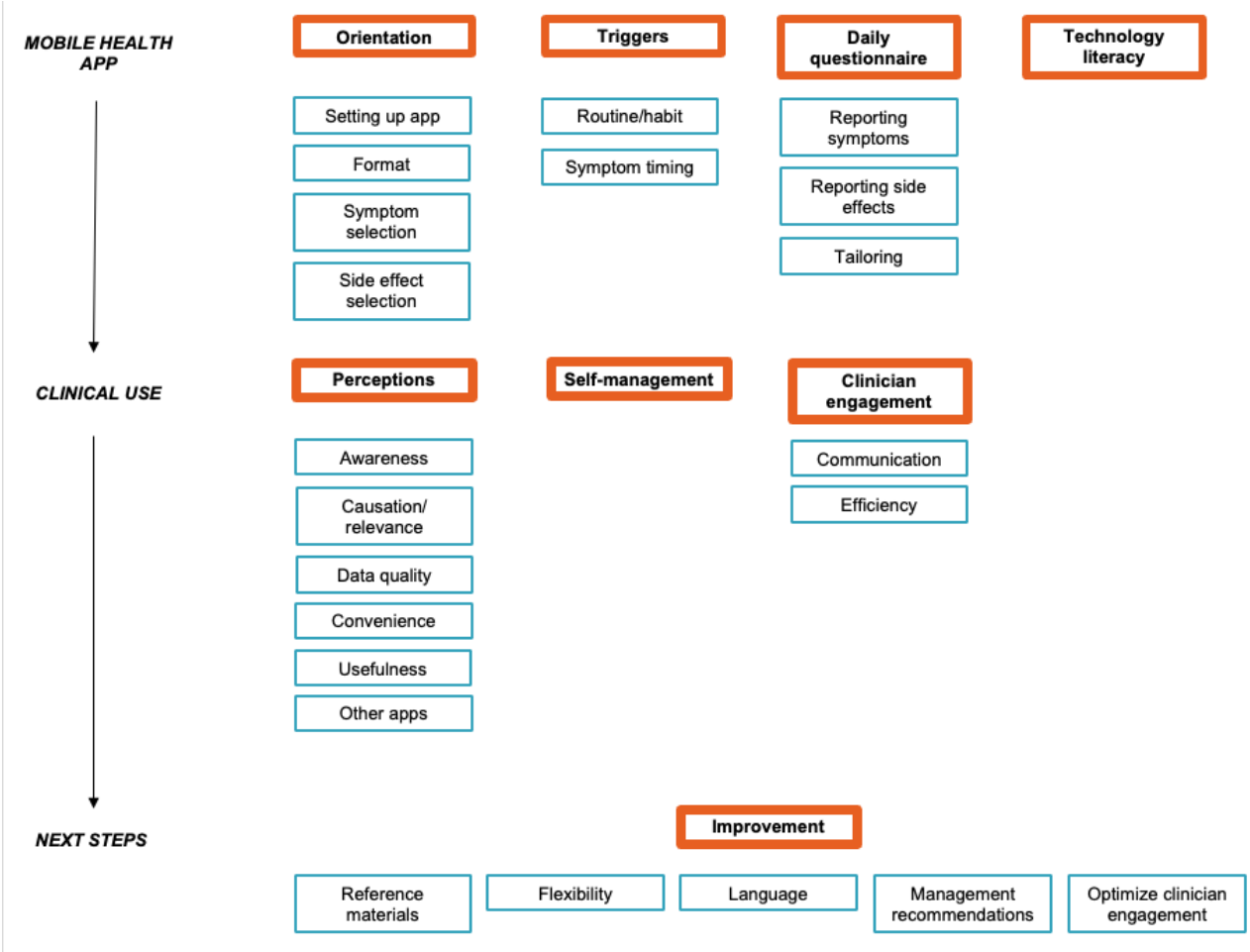


Table 2. Selected participant quotations organized by themes and codes.

Themes and codes	Quotations
Mobile health app	
Orientation : participants discussed the orientation process, from the assistance they received while setting up the app to the symptom and side-effect selection	
Setting up app	"Fairly easy. Good to have someone on the other line to take me through it step by step."
	"Instructions were very straightforward, and the graphics were easy to read."
Format	"The way it was done in the study was good: email and phone orientation."
	"Would prefer written documentation in case there were issues."
	"Send link by email and I can follow instructions—I don't need any human contact."
Symptom selection	"Disappointed that it was limited to three, would have done up to seven to eight."
	"I have more symptoms than the survey asked about—the top three weren't enough for me."
	"The more items I was keeping track of, the more I was attuned to my symptoms, so more questions would be good."
Side-effect selection	"Would have preferred to track more side effects. Initially thought some of these were just symptoms, not side effects, like runny nose. I also have allergies, but definitely could have been tamsulosin."
Triggers: participants discussed their experiences and preferences regarding reminders and notifications from the mobile app	
Routine or habit	"Never received notifications, I went in and answered by habit."
	"Just did the questions before bed, whether I got the notification or not yet that day. I was getting daily notifications."
Symptom timing	"Sometimes I would remember to answer questions in the morning before the notification—my symptoms were mostly during the night, so I remembered to log in the morning."
Daily questionnaire : participants discussed their experiences completing daily questionnaires within the mobile app	
Reporting symptoms	"Asked about symptoms that weren't applicable and did not ask about leakage."
	"Maybe some questions seemed like they would be too simple for others but were good for me."
Reporting side effects	"Impossible to answer—whatever side effects could be caused by other things. All of that was kind of irrelevant—couldn't relate anything to Flomax."
	"Hard to distinguish which symptoms are from what."
Tailoring	"Frustrated because some issues were not questioned in the survey. Not sure if I'm having them because of meds or not. The survey was simplistic, basic. Would have preferred more depth—and to be able to note the thoughts before this interview."
	"Some questions didn't apply."
Technology literacy: participants discussed their ability to engage with technology	
N/A[a]	"It's very easy for everybody, even people like me, old and not very good with technology."
Clinical use	
Perceptions : participants discussed their perceptions as a result of using the mobile app	
Awareness	"Brought more awareness to the issue; and now I think about it more than before. I kind of wonder how much it affects my sleep and it brings the awareness to the forefront."
	"Became more attuned to your symptoms, more attentive—had never previously paid attention."
	"Made me more aware of the topic."
	"Made me more conscious of [symptoms]. Haven't seen a doctor in a while, keeps the issues fresh on my mind so I can track them."
Causation or relevance	"No relationship between what Flomax does. The app tells you what you're feeling but doesn't say why you feel that way."
	"I didn't know if I was getting side effects, and I didn't know if they were related to the meds or just overall bladder problems; no baseline to compare to."[b]
	"What if you take eight meds? This is just one medication. Lots of meds have the same side effects, ie, headaches, dizziness."

Themes and codes	Quotations
Data quality	"The questions were limiting. The number of options for getting up at night were recorded as a range of numbers rather than a precise number. Trying to track things means you want to be more specific."
	"[Would prefer] having more longitudinal data—duration of symptoms."
	"Do you have this problem 0 times, 1-2, et cetera, well geez I'm more like 0-1, but there's no way to put 0-1. Having a selection process that takes in the range of options, I would suspect most people the frequency is variable from day to day. So, if you have 0-2, that certainly includes 0-1 and 1-2, and I would think that frequency question should be taken into account, maybe the lower levels more inclusively, rather than feeling like there's a gap or uncertainty in the choice."
Convenience	"Would have loved to continue this study, if possible, I could always record on paper but it's not as convenient. With the app, it's just clicking buttons."
	"It was short enough to stick to every day."
Usefulness	"I learned some things that I didn't know about my condition before; I was surprised about what I learned about my symptoms—'cause then I can talk to my doctor about it."
	"Same answers every day. It's a one-way survey, not a dialogue—not sure if the answers are good enough to change management."
	"Nice to have a tool to help with aging issues."
Other apps	"The PERSONAL[c] app is more user friendly, and you only have to provide simple responses."
	"This app was very basic, not even an app, just a questionnaire. An app tries to change behavior—there was nothing in the app that seems to have any influence on behavior."
Self-management: participants discuss using the questionnaires to guide self-management	
N/A	"More awareness of symptoms—the morning timing is very appropriate for answering the questions. So maybe that day based on my doctor's recommendations I will change my diet, for example, citrus/salt changes. I analyze why certain nights are worse than others."
N/A	"I'm going to talk to my doctor and tell him that I did this and tell him what I learned."
Clinician engagement: participants discuss their preferences for clinician engagement	
Communication	"I would like my doctor to have every scintilla of data they can have."
	"It made me more aware of leakage, because of incomplete emptying, and I want to share this with her."
Efficiency	"It would help: the more that you can do electronically to help your doctor, the less time they have to spend on the office visits, especially given how busy they are. That would be helpful in monitoring."
	"I think it would help that we would be looking at this issue a little more closely as opposed to a casual conversation during the visit. I think the doctor would be able to monitor the frequency of symptoms I'm having either throughout the night or the day and we can decide if the dosage needs to be changed."
Next steps	
Improvement: participants discuss potential areas of improvement for the mobile app and its integration into clinical care	
Reference materials	"Maybe save [the symptom list] for participants so they can reference which were symptoms versus side effects."
Flexibility	"More open ended, the questions from the app were overly simple."
	"Wanted afternoon/evening reminders, but was getting them in the morning, but I was reflecting on the day before rather than for the next 24 hours and that based on the person who set it up and I thought we could change it from AM to PM."
Language	"Based on experience as a nurse in Oakland where 90% of patients were Spanish speaking, many underserved patients over the age of 60 would not be able to benefit from this app as it currently stands."
	"Because English isn't first my language, something to read would be best so I can look it up on Google Translate."
Management recommendations	"There were no directions on how to manage one's symptoms based on the answers."
	"Maybe add questions in the app about how many cups of water to better help with management; could have scores/goals incorporated in the app to keep you on track, like Fitbit."
Optimize clinician engagement	"Yes—of note, I had a video appointment with my urologist during the 2 weeks, and told him about the study. My urologist's reaction was 'zero' because the urologist didn't know what it was."

[a]N/A: not applicable; there were no codes under this theme.

[b]N-of-1 is supposed to help establish a baseline to help reduce this type of confusion.

[c]PERSONAL: Placebo-Controlled, Randomized, Patient-Selected Outcomes, N-of-1 Trials.

Discussion

In sum, mHealth apps may play an important role in the chronic management of LUTS. In this study, the key themes that emerged from our qualitative data were (1) orientation, (2) triggers, (3) daily questionnaire, (4) technological literacy, (5) perceptions, (6) self-management, (7) clinician engagement, and (8) improvement. Consistent with prior literature [12,13], our participants had adequate technology literacy to use this app on a daily basis. The questionnaires were short enough for daily adherence, though a subset of participants with more severe or poorly controlled symptoms believed the questionnaires were not encompassing enough. These participants either felt that the app's closed-ended question format impeded accurate reporting, that they wanted to track symptoms that were not selected, or they wanted to track symptoms in addition to the ones selected. Given the desire for increased flexibility in communication, certain patients may benefit from a free-text or messaging option. Some patients taking multiple medications expressed frustration at not being able to differentiate which side effects were attributable to tamsulosin, further supporting the need for individualized tracking and potentially individual crossover trials (eg, N-of-1 studies). Patients with prior experience logging health information on a daily basis mentioned that mobile apps are a great way to store data and communicate important information with their clinicians, especially compared with recording on paper. Participants with more severe symptoms seemed less empowered to use their daily logs to change management, potentially because even their clinic visits with their clinicians failed to yield symptomatic improvement, among other reasons.

Most American men and adults over the age of 65 years own a smartphone [14]. Mobile apps are widely used by the general public, and the number of urology-related mobile apps has multiplied in recent decades [15]. These apps support patients with a wide range of urologic conditions, but their usage is unstandardized and unregulated [16]; in addition, mobile app usage remains inconsistently integrated with routine clinical care [17].

Studies support the need for expert involvement in mobile app development, dissemination, and regulation [15,16,18,19]. In our study, patients were unanimously willing to share this health data with their clinicians with the goal of optimizing their urologic care. Likewise, in a British study, urologists reported considerable interest in incorporating various mobile apps into their urologic practices [20]. Leveraging the doctor-patient relationship in the early phases of using the PERSONAL app would likely mitigate both patient confusion regarding which symptoms are medication side effects versus isolated symptoms, as well as clinician confusion when their patients ask to discuss data from a mobile app they are using to track their symptoms. Becoming more aware of symptoms via a mobile app may also help patients hone their questions when visiting their physician. Ultimately, mobile apps may become an important part of LUTS management, such as motivating patients and providers to stop chronic tamsulosin therapy if it is no longer helpful or causing nonspecific yet bothersome side effects. Tracking symptoms regularly using an mHealth app could also potentially help identify low-frequency, but serious, events associated with tamsulosin use, such as falls [21].

This study offers insights into the benefits and patient concerns associated with tracking symptoms via a mobile app among older men with benign urologic conditions, but there are some limitations. Traditionally, qualitative interviews are audio-recorded unstructured or semistructured interviews. Given the focus of the study, the homogeneous population, and the strong convergence of responses around a few common themes, we likely reached data saturation among 19 participants [22]. We solicited a wide range of responses, but less structured interviews in more diverse patient populations are warranted in future studies; additional qualitative studies involving clinical providers could inform the best ways to incorporate the use of mobile apps into clinical care. While qualitative interviews allow study participants to discuss their experiences in more depth, some participants may have felt less candid than they would answering an online form.

Despite the limitations, this is the first mixed methods study to examine the use of a mobile app to track symptoms in older men with LUTS. Importantly, these results suggest that mobile apps designed to improve symptom awareness and management may be adapted to benefit older men with LUTS due to BPH. As the use of mobile apps becomes increasingly popular, their usage in the health care setting will require further optimization. Tailoring to individual patients' health and technology literacy levels, language proficiency, and symptom severity will be critical to maximizing the efficacy of these digital interventions. Incorporating the use of apps into clinical practice may play another important role, pending future study.

Acknowledgments

We would like to acknowledge our study participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study inclusion and exclusion criteria.
[DOCX File , 15 KB - humanfactors_v8i4e30767_app1.docx]

Multimedia Appendix 2
Interview guide.
[DOCX File , 16 KB - humanfactors_v8i4e30767_app2.docx]

Multimedia Appendix 3
Participant feedback.
[DOCX File , 26 KB - humanfactors_v8i4e30767_app3.docx]

References

1. D'Ancona C, Haylen B, Oelke M, Abranches-Monteiro L, Arnold E, Goldman H, et al. The International Continence Society (ICS) report on the terminology for adult male lower urinary tract and pelvic floor symptoms and dysfunction. Neurourol Urodyn 2019 Feb;38(2):433-477. [doi: 10.1002/nau.23897] [Medline: 30681183]
2. Washington III SL, Shinohara K. Disorders of the bladder, prostate, and seminal vesicles. In: McAninch JW, Lue TF, editors. Smith & Tanagho's General Urology. 19th edition. New York, NY: McGraw Hill; 2020:585.
3. Coyne KS, Sexton CC, Thompson CL, Milsom I, Irwin D, Kopp ZS, et al. The prevalence of lower urinary tract symptoms (LUTS) in the USA, the UK and Sweden: Results from the Epidemiology of LUTS (EpiLUTS) study. BJU Int 2009 Aug;104(3):352-360. [doi: 10.1111/j.1464-410X.2009.08427.x] [Medline: 19281467]
4. van Leeuwen JHS, Castro R, Busse M, Bemelmans BL. The placebo effect in the pharmacologic treatment of patients with lower urinary tract symptoms. Eur Urol 2006 Sep;50(3):440-452; discussion 453. [doi: 10.1016/j.eururo.2006.05.014] [Medline: 16753253]
5. Malde S, Umbach R, Wheeler JR, Lytvyn L, Cornu JN, Gacci M, et al. A systematic review of patients' values, preferences, and expectations for the diagnosis and treatment of male lower urinary tract symptoms. Eur Urol 2021 Jun;79(6):796-809. [doi: 10.1016/j.eururo.2020.12.019] [Medline: 33461781]
6. Kim JH, Kwon S, Shim SR, Sun HY, Ko YM, Chun D, et al. Validation and reliability of a smartphone application for the International Prostate Symptom Score questionnaire: A randomized repeated measures crossover study. J Med Internet Res 2014 Feb 10;16(2):e38 [FREE Full text] [doi: 10.2196/jmir.3042] [Medline: 24513507]
7. Kwon W, Kim SH, Kim S, Joung JY, Chung J, Lee KH, et al. Changes in urination according to the sound of running water using a mobile phone application. PLoS One 2015;10(5):e0126798 [FREE Full text] [doi: 10.1371/journal.pone.0126798] [Medline: 25978378]
8. Choo MS, Jeong SJ, Cho SY, Yoo C, Jeong CW, Ku JH, et al. Development of decision support formulas for the prediction of bladder outlet obstruction and prostatic surgery in patients with lower urinary tract symptom/benign prostatic hyperplasia: Part II, External validation and usability testing of a smartphone app. Int Neurourol J 2017 Apr;21(Suppl 1):S66-S75 [FREE Full text] [doi: 10.5213/inj.1734854.427] [Medline: 28446011]
9. Bauer SR, Breyer BN, Oni-Orisan A, Steinman MA, Sim I, McCulloch CE, et al. PERSONAL: Feasibility study protocol for placebo-controlled, randomized n-of-1 trials of tamsulosin for lower urinary tract symptoms. Front Digit Health 2020;2:7 [FREE Full text] [doi: 10.3389/fdgth.2020.00007] [Medline: 34713020]
10. Creswell JW, Fetters MD, Ivankova NV. Designing a mixed methods study in primary care. Ann Fam Med 2004;2(1):7-12 [FREE Full text] [doi: 10.1370/afm.104] [Medline: 15053277]
11. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): A 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
12. Morselli S, Liaci A, Nicoletti R, Pecoraro A, Gemma L, Polverino P, et al. The use of a novel smartphone app for monitoring male LUTS treatment during the COVID-19 outbreak. Prostate Cancer Prostatic Dis 2020 Dec;23(4):724-726 [FREE Full text] [doi: 10.1038/s41391-020-0253-z] [Medline: 32665609]
13. Anthony Berauk VL, Murugiah MK, Soh YC, Chuan Sheng Y, Wong TW, Ming LC. Mobile health applications for caring of older people: Review and comparison. Ther Innov Regul Sci 2018 May;52(3):374-382. [doi: 10.1177/2168479017725556] [Medline: 29714532]
14. Mobile fact sheet. Pew Research Center. 2019 Jun 12. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2021-12-16]
15. Mantica G, Malinaric R, Dotta F, Paraboschi I, Guano G, Rebuffo S, et al. Urology apps: Overview of current types and use. Cent European J Urol 2020;73(3):369-372 [FREE Full text] [doi: 10.5173/ceju.2020.0090] [Medline: 33133667]
16. Akbar S, Coiera E, Magrabi F. Safety concerns with consumer-facing mobile health applications and their consequences: A scoping review. J Am Med Inform Assoc 2020 Feb 01;27(2):330-340 [FREE Full text] [doi: 10.1093/jamia/ocz175] [Medline: 31599936]
17. Tapiero S, Yoon R, Jefferson F, Sung J, Limfueco L, Cottone C, et al. Smartphone technology and its applications in urology: A review of the literature. World J Urol 2020 Oct;38(10):2393-2410. [doi: 10.1007/s00345-019-02960-y] [Medline: 31598754]

18. Pereira-Azevedo N, Carrasquinho E, Cardoso de Oliveira E, Cavadas V, Osório L, Fraga A, et al. mHealth in urology: A review of experts' involvement in app development. PLoS One 2015;10(5):e0125547 [FREE Full text] [doi: 10.1371/journal.pone.0125547] [Medline: 25984916]
19. Pereira-Azevedo N, Osório L, Cavadas V, Fraga A, Carrasquinho E, Cardoso de Oliveira E, et al. Expert involvement predicts mHealth app downloads: Multivariate regression analysis of urology apps. JMIR Mhealth Uhealth 2016 Jul 15;4(3):e86 [FREE Full text] [doi: 10.2196/mhealth.5738] [Medline: 27421338]
20. Dempster NJ, Risk R, Clark R, Meddings RN. Urologists' usage and perceptions of urological apps. J Telemed Telecare 2014 Dec;20(8):450-453. [doi: 10.1177/1357633X14555622] [Medline: 25316043]
21. Welk B, McArthur E, Fraser L, Hayward J, Dixon S, Hwang YJ, et al. The risk of fall and fracture with the initiation of a prostate-selective α antagonist: A population based cohort study. BMJ 2015 Oct 26;351:h5398 [FREE Full text] [doi: 10.1136/bmj.h5398] [Medline: 26502947]
22. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: Systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):148 [FREE Full text] [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]

Abbreviations

BPH: benign prostatic hyperplasia
COREQ: Consolidated Criteria for Reporting Qualitative Research
LUTS: lower urinary tract symptoms
mHealth: mobile health
PERSONAL: Placebo-Controlled, Randomized, Patient-Selected Outcomes, N-of-1 Trials
REDCap: Research Electronic Data Capture

Edited by A Kushniruk; submitted 28.05.21; peer-reviewed by T Johnson II, V Cavadas, A Hidki; comments to author 04.07.21; revised version received 28.08.21; accepted 31.08.21; published 24.12.21.

Please cite as:
Wang EY, Breyer BN, Lee AW, Rios N, Oni-Orisan A, Steinman MA, Sim I, Kenfield SA, Bauer SR
Perceptions of Older Men Using a Mobile Health App to Monitor Lower Urinary Tract Symptoms and Tamsulosin Side Effects: Mixed Methods Study
JMIR Hum Factors 2021;8(4):e30767
URL: https://humanfactors.jmir.org/2021/4/e30767
doi:10.2196/30767
PMID:34951599

Original Paper

Usability of an At-Home Anterior Nares SARS-CoV-2 RT-PCR Sample Collection Kit: Human Factors Feasibility Study

Laura E Strong[1], PhD; Irene Middendorf[1], MS; Michelle Turner[1], MS; David K Edwards V[1], PhD; Varun Sama[1], MS; Joshua Mou[1], MD; K Colleen Adams[1], MTSC

Exact Sciences Corporation, Madison, WI, United States

Corresponding Author:
K Colleen Adams, MTSC
Exact Sciences Corporation
5505 Endeavor Lane
Madison, WI, 53719
United States
Phone: 1 6088006531
Email: cadams@exactsciences.com

Abstract

Background: Readily available testing for SARS-CoV-2 is necessary to mitigate COVID-19 disease outbreaks. At-home collection kits, in which samples are self-collected without requiring a laboratory or clinic visit and sent to an external laboratory for testing, can provide convenient testing to those with barriers to access. They can prevent unnecessary exposure between patient and clinical staff, increase access for patients with disabilities or remote workers, and decrease burdens on health care resources, such as provider time and personal protective equipment. Exact Sciences developed an at-home collection kit for samples to be tested to detect SARS-CoV-2 that includes an Instructions for Use (IFU) document, which guides people without prior experience on collecting a nasal swab sample. Demonstrating successful sample collection and usability is critical to ensure that these samples meet the same high-quality sample collection standards as samples collected in clinics.

Objective: The aim of this study was to determine the usability of a SARS-CoV-2 at-home nasal swab sample collection kit.

Methods: A human factors usability study was conducted with 30 subjects without prior medical, laboratory, or health care training and without COVID-19 sample self-collection experience. Subjects were observed while they followed the IFU for the at-home sample collection portion of the SARS-CoV-2 test in a setting that simulated a home environment. IFU usability was further evaluated by requiring the subjects to complete a survey, answer comprehension questions, provide written feedback, and respond to questions from the observer about problems during use.

Results: All 30 subjects successfully completed the sample collection process, and all 30 samples were determined by reverse transcription–polymerase chain reaction (RT-PCR) testing to meet quality standards for SARS-CoV-2 testing. The subjects' written feedback and comments revealed several recommendations to improve the IFU.

Conclusions: The study demonstrated the overall usability of an at-home SARS-CoV-2 collection kit. Various feedback mechanisms provided opportunities to improve the wording and graphics for some critical tasks, including placing the label correctly on the tube. A modified IFU was prepared based on study outcomes.

(JMIR Hum Factors 2021;8(4):e29234) doi:10.2196/29234

KEYWORDS

COVID-19 testing; at-home collection kit; SARS-CoV-2; feasibility studies; self-collection; usability study; COVID-19

Introduction

The global pandemic caused by SARS-CoV-2 has resulted in 223 million confirmed cases of COVID-19, including 4.6 million deaths, as of September 2021, according to the World Health Organization [1]. In the United States, there have been more than 40.3 million reported COVID-19 cases and more than 649,000 deaths as of September 2021 [1]. The transmission of COVID-19 has been shown to be contained with a combination of isolation practices, including wearing masks, physical distancing, and lockdown measures [2], and widespread immunization with effective COVID-19 vaccines [3]. In the United States, more than 175 million people have been fully vaccinated as of September 2021 using one of three COVID-19

vaccines [4-6] currently authorized for emergency use by the US Food and Drug Administration (FDA)—in August 2021, the FDA approved Comirnaty, known previously as the Pfizer-BioNTech COVID-19 vaccine, for individuals 16 years of age or older [7].

Despite this monumental progress, significant challenges remain to manage the ongoing pandemic. Many eligible Americans have not been vaccinated [1], and the emergence of increasingly virulent strains, including the Delta variant [8], have resulted in increased hospitalizations and deaths throughout the United States [9,10]. Widespread testing can help public health officials to better monitor the progression of the pandemic, identify emerging variants, and identify individuals with COVID-19, particularly those with asymptomatic disease.

One approach to broadening access to SARS-CoV-2 testing has been the development of at-home sample collection kits that could be used safely and effectively by people without medical or laboratory experience. Samples can be collected without needing to travel to a medical center, and the samples can be shipped and later processed at a laboratory or health care facility or tested at home [11]. At-home sample collection offers multiple advantages to combat the COVID-19 pandemic: it can prevent unnecessary exposure between patients and clinical staff during collection; improve access for elderly patients, patients with disabilities, or remote workers; reduce the need for personal protective equipment; and shift the logistics of collection from overburdened clinical sites to commercial delivery services.

Surveys on the perception of at-home COVID-19 sample collection and tests have demonstrated a broad willingness to complete such collection and confidence in the sample preparation and quality [12]. Multiple studies have been conducted recently to compare SARS-CoV-2 test results from self-collected samples to those collected by health care workers [13-18]. Recently, a large-scale population-based study on the applicability of COVID-19 self-testing demonstrated that most participants collected the sample correctly the first time, and that test results showed comparable performance to those collected by health care professionals [19].

To investigate the usability of an at-home collection process, usability studies should be conducted to ensure that these samples meet the same quality standards as clinician-collected samples. Here, we describe the results of a human factors usability study, conducted early in the pandemic, for the at-home sample collection kit, herein referred to as the "SARS-CoV-2 at-home collection kit," for use with the SARS-CoV-2 (N gene detection) Test, both of which were developed by Exact Sciences. Exact Sciences is a molecular diagnostics company that manufactures an at-home screening test for colorectal cancer and developed the SARS-CoV-2 at-home collection kit and test in response to the global pandemic. Additional details related to the SARS-CoV-2 test are available in the FDA's Emergency Use Authorization (EUA) documentation in Multimedia Appendix 1.

The goal of this study was to determine the usability of the SARS-CoV-2 at-home collection kit, and our primary endpoint was the percentage of samples collected from study participants that returned a valid SARS-CoV-2 result.

This study was conducted in May 2020, during the early months of the pandemic when very little information about COVID-19 pathogenesis was available, and the protocol was designed based on standards for human factors usability study design [20]. By publishing our methodology and outcomes, we hope to provide a blueprint for future studies to ensure that the usability of other at-home collection kits can be quickly evaluated during public health crises or similar situations where urgency is required.

Methods

SARS-CoV-2 Test

The SARS-CoV-2 (N gene detection) Test was developed by Exact Sciences and received EUA from the FDA on May 22, 2020, via EUA200367 (Multimedia Appendix 1). A summary of the SARS-CoV-2 test characteristics is provided in Table 1. This is a reverse transcription–polymerase chain reaction (RT-PCR)–based test that evaluates upper respiratory samples, including those collected with an anterior nares (ie, nasal) swab, to detect regions within the nucleocapsid (N) gene of the novel coronavirus (nCoV), specifically the nCoV_N1 and nCoV_N2 regions. Human ribonuclease P (RNase P), a gene expressed ubiquitously in human cells regardless of COVID-19 infection, serves as a control to demonstrate that usable samples were collected and provided to the lab, and that all testing processes were successfully completed. In validation studies, the test demonstrated no cross-reactivity with a panel of known respiratory pathogens. Its preclinical test performance in a collection of test samples showed positive percent agreement of 95% (38 out of 40 samples) and negative percent agreement of 100% (38 out of 38 samples) with another FDA-authorized SARS-CoV-2 RT-PCR-based test.

The SARS-CoV-2 at-home collection kit contains the following: a sterile, individually wrapped nasal swab with a polyester tip with plastic handle; a 2-mL transport tube containing 0.9% saline; an Instructions for Use (IFU) document; a biohazard bag; an absorbent pad; a specimen identification label; and a UN3373-labeled shipping container.

Table 1. Performance of the SARS-CoV-2 detection test from Exact Sciences.

Test[a] characteristic	Details
Test name	SARS-CoV-2 (N gene detection) Test
Type of test	Real-time RT-PCR[b]
Gene regions detected	
SARS-CoV-2	nCoV[c]_N1 and nCoV_N2 regions of the nucleocapsid (N) gene
Control	Ribonuclease P human gene locus
Limit of detection	2.6 genome copies/µL sample
Cross-reactivity	13 other respiratory pathogens not detected[d]
Preclinical test performance[e] (n=78 samples)	
Positive percent agreement (38 out of 40 samples), % (95% CI)	95.0 (83.5-98.6)
Negative percent agreement (38 out of 38 samples), % (95% CI)	100 (90.8-100)

[a]Test details were obtained from the US Food and Drug Administration (FDA) Emergency Use Authorization (EUA) summary (Multimedia Appendix 1).

[b]RT-PCR: reverse transcription–polymerase chain reaction.

[c]nCoV: novel coronavirus.

[d]The detection assay was conducted using NATtrol Respiratory Pathogen Panel-1 (NATRPP-1) from Zeptometrix.

[e]In comparison with another COVID-19 RT-PCR test with FDA EUA.

Study Objectives and Subjects

The main objective was to determine the usability of the SARS-CoV-2 at-home collection kit for the collection and mailing of a nasal swab sample to the testing laboratory. The primary endpoint was the percentage of samples from the study participants that returned a valid SARS-CoV-2 test result, either positive or negative, both of which require a detectable level of RNase P. The target percentage was 80%, given that the subjects were minimally trained and were inexperienced in sample self-collection. There were five secondary objectives: (1) evaluate the perceived usability of the IFU, (2) evaluate the comprehension of the IFU by the subject, (3) identify problems that occur while following the IFU, (4) evaluate the root causes of problems that occur while following the IFU, and (5) develop strategies to mitigate problems occurring while following the IFU. A total of 30 patients from a workforce population that met established inclusion and exclusion criteria were enrolled in this study. Steps were taken to recruit subjects of varying ages and educational statuses, which included a manual review of participants by the study team.

The study was conducted in accordance with state and federal regulatory requirements, as well as the general principles set forth in the International Ethical Guidelines for Biomedical Research Involving Human Subjects [21] and the Declaration of Helsinki [22]. The study was approved by the Western Institutional Review Board (WIRB)–Copernicus Group Institutional Review Board (No. 20201763) and all subjects provided written informed consent prior to study participation.

The inclusion criterion was the ability to provide informed consent, and the exclusion criteria were prior medical or laboratory training, prior experience with COVID-19 sample self-collection, and prior SARS-CoV-2 testing. For each enrolled subject, usability of the IFU was determined based on successful completion of the self-collection of a nasal swab sample, which included a valid SARS-CoV-2 test result. All subjects completed the study.

Study Design

A use-related Failure Mode Effects Analysis approach was used to determine potential hazards and their associated risks during use of the SARS-CoV-2 at-home collection kit. Subjects completed a survey form with demographic information, including race and ethnicity [23], highest education level obtained, and prior experience with medical or laboratory training and COVID-19 sample self-collection. Subjects who provided informed consent were provided an overview of the clinical study procedures, including guidance that they would be observed during the sample collection and answer questions related to their experience during the sample collection. The study consisted of two parts: simulated use, in which sample collection was simulated by a subject while monitored by an observer; and postsimulation evaluation, in which the subject completed survey questions and provided feedback on the collection process. The overall study design and methodology is summarized in Figure 1.

The sample collection took place in a simulated home environment in a conference room with a table that served as a large surface area, similar to a countertop found in a kitchen or bathroom. The room included common household items, such as hand sanitizer, pens, pencils, paper towels, and a wastebasket. Since a sink was not available in the room, a large bowl labelled “SINK” was provided next to the hand sanitizer to simulate a sink for handwashing with soap and water.

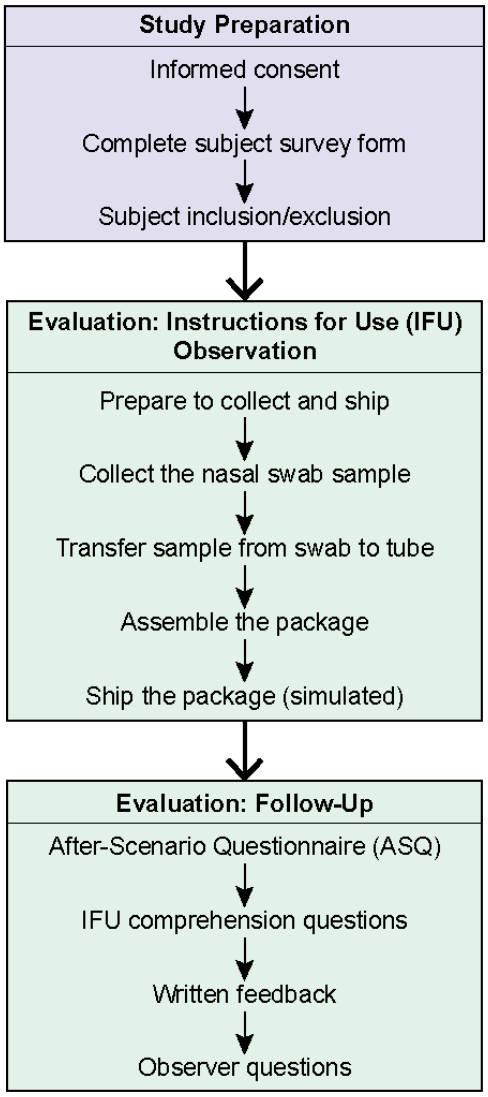
Before beginning the sample collection, the observer oriented the subject to the simulated environment, making them aware of items available to them, without indicating that these would be required for the sample collection to reduce bias to the

subject. To begin the sample collection, the observer instructed the subject to retrieve an available kit and to begin the collection; subjects were then observed while following the IFU for the SARS-CoV-2 at-home collection kit. The identification and classification of tasks were made by the study researchers (see Figure 1 for complete list of tasks), and the study participants were blinded to the task categories.

Following completion of all steps in the IFU, the subject was then instructed to place the sample package in a designated area within the conference room. The subject then provided feedback on the usability of the IFU by completing the After-Scenario Questionnaire (ASQ) [24-26], answered comprehension questions, provided written feedback on the experience, and addressed questions from the observer about observed problems during use.

The samples collected by the subjects were tested for SARS-CoV-2 by Exact Sciences Laboratories. Subjects remained blinded to the test results.

Figure 1. Overview of human factors usability study procedures.



Results

The study was conducted using the SARS-CoV-2 at-home collection kit from Exact Sciences; test characteristics are summarized in Table 1 (Multimedia Appendix 1). Briefly, the laboratory test used RT-PCR to detect SARS-CoV-2 from a sample collected using an anterior nares (ie, nasal) swab, using detection of human RNase P as a control.

For the human factors usability study, 30 subjects were enrolled to simulate at-home sample collection and provide feedback during the follow-up evaluation (Figure 1). The characteristics of the study subjects (N=30) are described in Table 2. The mean age of the subjects was 38.0 (SD 9.7) years, and no subjects were older than 65 years. Most subjects were White (n=26, 87%) and non-Hispanic or non-Latino (n=25, 83%), and 77% (n=23) of subjects had more than a high school education. After completing the usability study, all 30 subjects' self-collected samples resulted in a valid SARS-CoV-2 test result, and all were negative for SARS-CoV-2. Sample validity was determined by successful detection of human RNase P.

Subjects completed the simulated sample collection according to an IFU document that described the 26 tasks required to prepare, collect, and ship a nasal swab sample (see Multimedia Appendix 2 for complete list of tasks). These tasks were divided into "critical tasks," in which use errors or failure to complete

would have a negative clinical impact, such as invalid or delayed test results, and "essential tasks," which were important for test completion but did not pose an immediate risk to the sample. Out of the 26 tasks, 15 were categorized as "critical" and were the primary focus for evaluating and improving the IFU based on study outcomes and subject feedback.

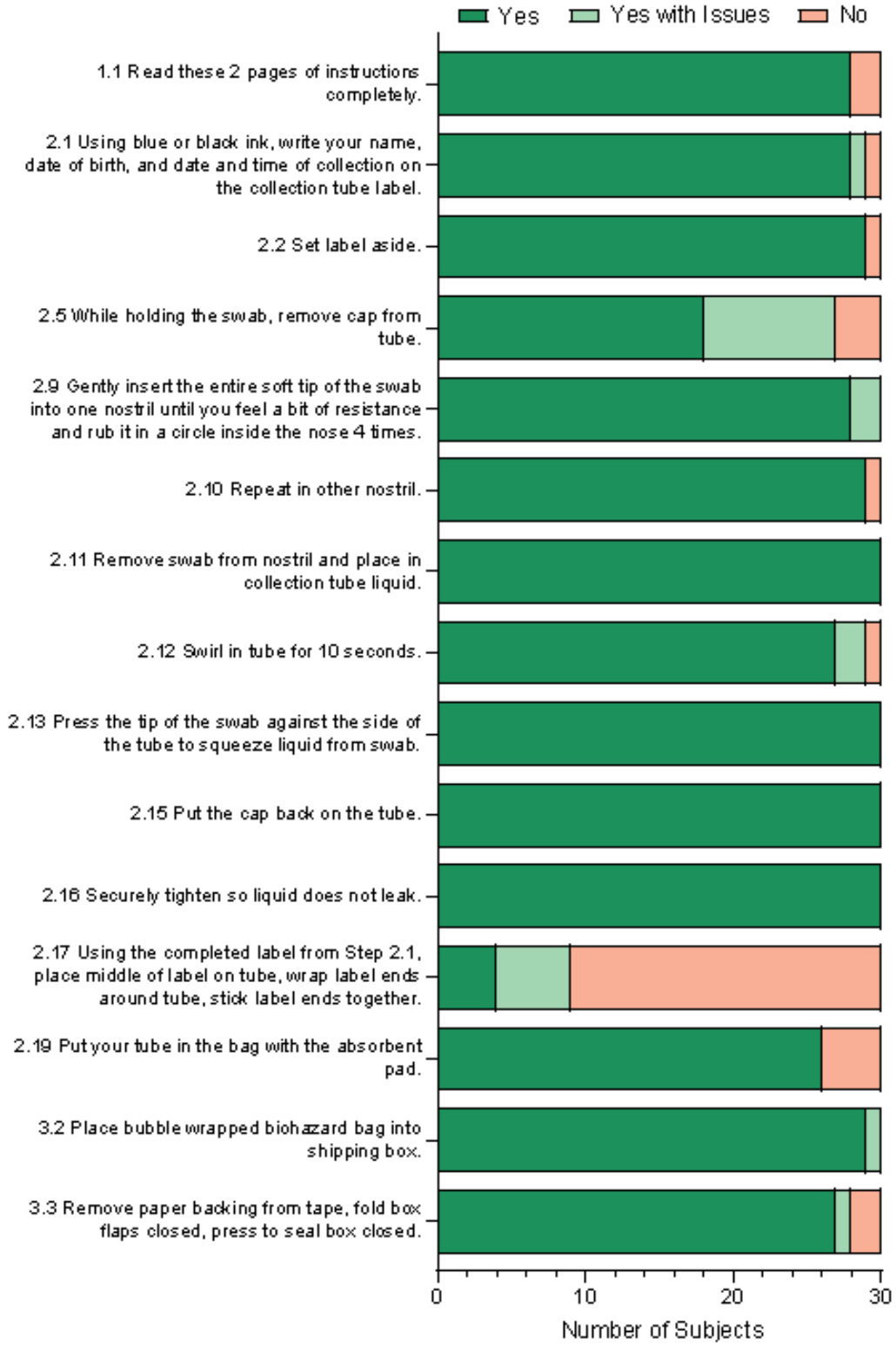
Overall, 14 out of 15 critical tasks from the IFU were successfully completed by more than 80% of the subjects during the simulated sample collection (Figure 2). The task that the subjects encountered the greatest difficulty with was placing the label on the tube, which was not completed properly by 70% (21/30) of the subjects.

Table 2. Subject characteristics for human factors usability study.

Characteristic	Value (N=30)
Age (years), mean (SD)	38.0 (9.7)
Age category (years), n (%)	
<18	0 (0)
18-30	9 (30)
31-45	12 (40)
46-65	9 (30)
>65	0 (0)
Sex, n (%)	
Female	18 (60)
Male	12 (40)
Ethnicity, n (%)	
Hispanic or Latino	4 (13)
Non-Hispanic or non-Latino	25 (83)
Unknown	1 (3)
Race[a], n (%)	
American Indian or Alaska Native	0 (0)
Asian	2 (7)
Black or African American	2 (7)
Native Hawaiian or other Pacific Islander	0 (0)
White	26 (87)
Unknown	0 (0)
Education level, n (%)	
No high school	0 (0)
Some high school	0 (0)
High school degree only	7 (23)
College degree	18 (60)
Advanced degree	5 (17)

[a]Subjects had the option to report one or more categories for race; each participant selected an option for both ethnicity and race.

Figure 2. Observation of success in completing critical tasks using the SARS-CoV-2 at-home collection kit.



To evaluate the opinions of the subjects after the simulated sample collection, subjects completed the ASQ (Table 3) [26]. Responses were indicated on a scale of 1 to 7, with lower scores corresponding to higher satisfaction; scores of less than 3 indicated that subjects felt satisfied using the IFU for sample collection. For all questions among the 30 subjects, the mean overall ASQ score was 2.1 (SD 1.6), indicating overall satisfaction.

The subjects' written feedback and comments to the observer revealed several areas of potential improvement to the IFU. Most comments focused on references to the front and back of the IFU, handwashing, handling the absorbent pad, how far into the nostrils the nasal swab should be inserted (eg, use of the word "resistance"), and issues related to the tube label (eg, writing on the label and attaching it to the tube).

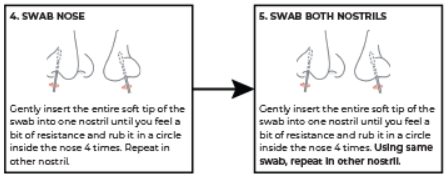
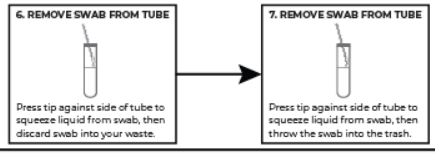
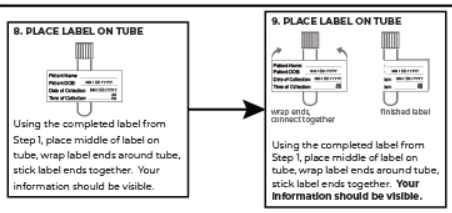
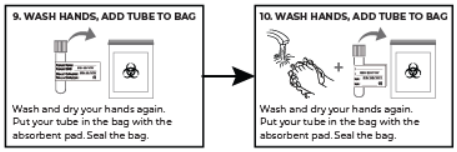
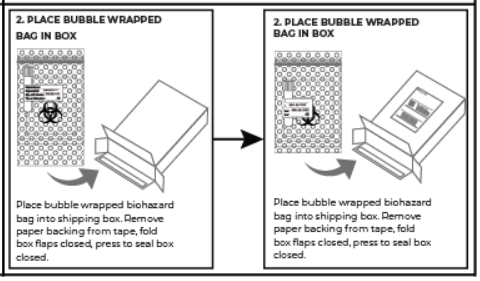
Results from the multiple sources of feedback collected during the feasibility study (Table 3) were combined to determine how

to improve the usability of the IFU (see Figure 3 for complete list of changes and corresponding rationale). This feedback helped improve the language and graphics describing how to place the label on the tube, which is the critical task with the lowest successful completion rate. Furthermore, subjects suggested minor wording changes to improve comprehension (eg, replacing "discard swab into your waste" with "throw swab into the trash"). The updated IFU is shown in Multimedia Appendix 3.

Table 3. Methods of evaluating Instructions for Use (IFU) sample collection and shipping tasks for SARS-CoV-2 at-home collection kit by the observer. After-Scenario Questionnaire (ASQ) questions were from Lewis [26].

Feedback	Categories or questions	Measurement
Observer evaluation of IFU	• Reading the instructions (1 step) • Preparing for collection (2 steps) • Preparing the tube label (2 steps) • Opening nasal swab (2 steps) • Removing the tube cap (4 steps) • Swabbing nose (2 steps) • Adding swab to tube (2 steps) • Removing swab from tube (2 steps) • Replacing the tube cap (2 steps) • Placing label on tube (1 step) • Washing hands and adding tube to bag (3 steps) • Placing bag in bubble wrap (1 step) • Placing bubble-wrapped bag in box (2 steps)	The observer selected one of four options: • Subject completed task with no issues • Subject completed task with issues or unexpected effort • Subject did not complete task or required assistance • Not applicable (subject discontinued participation).
ASQ questions	• ASQ1. Overall, I am satisfied with the ease of completing the tasks in this scenario. • ASQ2. Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario. • ASQ3. Overall, I am satisfied with the support information (online help, messages, and documentation) when completing the tasks.	The subject recorded their response on a numerical 7-point scale, ranging from 1 ("strongly agree") to 7 ("strongly disagree"), and "N/A" (not applicable) outside the scale.
IFU comprehension	• After collecting a nasal swab sample, when should a person ship it to the lab? • How should a person store the package with the nasal swab sample inside before shipping it back to the lab? • What could happen to your nasal swab sample if you do not follow the steps in the instructions for use?	The observer recorded the response as "correct" or "incorrect" (with the option to record free-form text and ask follow-up questions).
Written feedback	• What information in the IFU is confusing? • Is there anything we could do to make it easier to collect a nasal swab sample using these materials?	The observer recorded the response as free-form text (with the option to ask follow-up questions).
Observer questions	• Did the subject experience or report adverse events? • Were any protocol deviations noted? • Did the subject complete the study?	The observer recorded the response as "yes" or "no."

Figure 3. Revised Instructions for Use based on human factors usability study results. N/A: not applicable.

Critical Step Category	Original Instruction	Action Taken and Rationale	Updated Instruction
1.1	**READ THESE 2 PAGES OF INSTRUCTIONS COMPLETELY**	No change	N/A
2.1, 2.2	**PREPARE THE TUBE LABEL.** Using blue or black ink, write your name, date of birth, and date and time of collection on the collection tube label. Set label aside.	No change	N/A
2.5, 2.6, 2.7, 2.8	**REMOVE TUBE CAP.** While holding the swab, remove cap from tube. Be careful to not spill liquid. Do not drink liquid. Set cap aside.	No change	N/A
2.9, 2.10	**SWAB NOSE.** Gently insert the entire soft tip of the swab into one nostril until you feel a bit of resistance and rub it in a circle inside the nose 4 times. Repeat in other nostril.	Revised text and bolded to clearly indicate swabbing both nostrils with same swab based on participant feedback.	
2.11, 2.12	**ADD SWAB TO TUBE.** Remove swab from nostril and place in collection tube liquid. Swirl in tube for 10 seconds.	No change	N/A
2.13, 2.14	**REMOVE SWAB FROM TUBE.** Press tip against side of tube to squeeze liquid from swab, then discard swab into your waste.	Updated language from "discard swab into your waste" to "throw the swab into the trash" based on participant feedback.	
2.15, 2.16	**REPLACE TUBE CAP.** Put the cap back on the tube. Securely tighten so liquid does not leak.	No change	N/A
2.17	**PLACE LABEL ON TUBE.** Using the completed label from Step 2.1, place middle of label on tube, wrap label ends around tube, stick label ends together. Your information should be visible.	Updated graphic to better represent the correct label application based on user feedback. Added "wrap ends, stick together" text and arrows to indicate wrapping label ends around the tube. Added second image to represent the "finished label". Bolded last sentence to emphasize label placement result.	
2.18 – 2.20	**WASH HANDS, ADD TUBE TO BAG.** Wash and dry your hands again. Put your tube in the bag with the absorbent pad. Seal the bag.	Added handwashing image to emphasize instruction based on user feedback. Updated tube and label image to match previous instructions.	
3.2 – 3.3	**PLACE BUBBLE WRAPPED BAG IN BOX.** Place bubble wrapped biohazard bag into shipping box. Remove paper backing from tape, fold box flaps closed, press to seal box closed.	Updated tube and label image to match previous instructions.	

Discussion

Overall, this simulated at-home self-collection usability study was successful in that all 30 subjects collected samples that resulted in valid test results (100% success rate, exceeding the targeted 80%). Moreover, the ASQ scores were low, indicating acceptable agreement and satisfaction, and the written feedback and comments from subjects were combined with simulation data to improve the IFU for future patients undergoing COVID-19 testing using at-home specimen collection.

To evaluate sample quality, the presence of human RNase P, a gene expressed ubiquitously in human cells regardless of COVID-19 infection, provided a universal measurement of quality control. If the sample contained a detectable level of RNase P, then it was determined that the sample had sufficient RNA to be tested for the presence of SARS-CoV-2. Notably, this standard can be used for any sample obtained using the at-home SARS-CoV-2 collection kit, regardless of the method of collection or the positive or negative outcome of the test. The preclinical test characteristics of the SARS-CoV-2 test, summarized in Table 1, demonstrated high positive and negative percent agreement among samples of sufficient quality.

The study population was well distributed with respect to age and gender. The proportion with a college degree or higher (77%) was slightly higher than the local population in Madison, Wisconsin (58%), although race and ethnicity populations were

similar (ie, the White, non-Hispanic or non-Latino population in Madison is 74%) [27]. Importantly, the qualifications for study participants impacted the inclusion criteria, which required the exclusion of anyone with any scientific or laboratory experience.

In general, the ability to provide at-home sample collection to detect respiratory viruses could significantly improve the effectiveness of public health strategies in preventing the spread of disease during a pandemic. For the COVID-19 pandemic, at-home sample collection could (1) improve the ability to identify individuals with detectable SARS-CoV-2 RNA without the need to expose health care workers during testing or the public during travel to and from the testing site, (2) provide an alternative and likely more accessible testing workflow for patients, and (3) enable the epidemiological study of the natural history of disease without undue risk to the population.

This study had several limitations. Some limitations were the simulated nature of the home environment, the lack of access to shipment methods for subjects, and the lack of subjects over 65 years, which was a result of the workforce population recruited for the study. Other limitations, including the relatively small sample size, were based on limited access of materials and a prioritization to make this collection kit available as soon as possible due to the ongoing public health crisis. Changes driven by logistics or product considerations and typographical errors are included in the updated IFU but are beyond the scope of this publication. The strengths of this study were that all subjects were able to successfully follow the IFU to collect usable samples, the consistency of the completion of medium-risk tasks, and the constructive feedback on low-risk tasks that led to IFU improvements.

In conclusion, this study demonstrated the overall usability of the SARS-CoV-2 at-home collection kit, and feedback from the study was used to generate improved instructions for use. Overall, it provides additional information that at-home collection of specimens for use with COVID-19 tests can be conducted effectively by subjects without prior sample self-collection experience.

Acknowledgments

We would like to thank the following contributors, who were all at Exact Sciences during the study design and execution: Steve Shak, MD, for scientific advice; Erin Bork, MS, Hannah McCracken, Marina Pavlova, Corey Yackel, Sarah Dunlap, PhD, and Claire Anderson for data processing assistance; Kourtney Frick for graphic design support; Rob Fisher, PhD, for Instructions for Use design and feedback incorporation; and everyone at Exact Sciences whose efforts and assistance helped make this study successful. This study was funded by Exact Sciences Corporation.

Authors' Contributions

LES, MT, and KCA were responsible for conceptualization of the study. LES, IM, MT, and KCA were responsible for the study methodology. VS was responsible for software programming and data analysis. LES, MT, and VS were responsible for the formal analysis. IM was responsible for conducting the investigation and for project administration. DKEV was responsible for writing the original draft of the manuscript and for visualization of the results. LES, IM, MT, VS, DKEV, JM, and KCA were responsible for writing, reviewing, and editing the manuscript. LES, IM, and MT were responsible for project supervision.

Conflicts of Interest

MT, VS, DKEV, JM, and KCA are employees of Exact Sciences Corporation. LES and IM were employees of Exact Sciences Corporation when this study was conducted.

Multimedia Appendix 1
US Food and Drug Administration Emergency Use Authorization documentation for the Exact Sciences SARS-CoV-2 test.
[PDF File (Adobe PDF File), 444 KB - humanfactors_v8i4e29234_app1.pdf]

Multimedia Appendix 2
List of steps in Instructions for Use for the SARS-CoV-2 at-home collection kit. Critical steps are indicated in bold, which were not uniquely highlighted or bolded to subjects during the study.
[PDF File (Adobe PDF File), 466 KB - humanfactors_v8i4e29234_app2.pdf]

Multimedia Appendix 3
Final Instructions for Use document.
[PDF File (Adobe PDF File), 1069 KB - humanfactors_v8i4e29234_app3.pdf]

References

1. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. 2021. URL: https://covid19.who.int/ [accessed 2021-09-02]

2. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. Lancet 2020 Jun 27;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
3. Hodgson SH, Mansatta K, Mallett G, Harris V, Emary KRW, Pollard AJ. What defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2. Lancet Infect Dis 2021 Feb;21(2):e26-e35 [FREE Full text] [doi: 10.1016/S1473-3099(20)30773-8] [Medline: 33125914]
4. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020 Dec 31;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]
5. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021 Feb 04;384(5):403-416 [FREE Full text] [doi: 10.1056/NEJMoa2035389] [Medline: 33378609]
6. Sadoff J, Gray G, Vandebosch A, Cárdenas V, Shukarev G, Grinsztejn B, ENSEMBLE Study Group. Safety and efficacy of single-dose Ad26.COV2.S vaccine against Covid-19. N Engl J Med 2021 Jun 10;384(23):2187-2201 [FREE Full text] [doi: 10.1056/NEJMoa2101544] [Medline: 33882225]
7. FDA approves first COVID-19 vaccine: Approval signifies key achievement for public health. US Food and Drug Administration. 2021 Aug 23. URL: https://www.fda.gov/news-events/press-announcements/fda-approves-first-covid-19-vaccine [accessed 2021-09-01]
8. Gómez CE, Perdiguero B, Esteban M. Emerging SARS-CoV-2 variants and impact in global vaccination programs against SARS-CoV-2/COVID-19. Vaccines (Basel) 2021 Mar 11;9(3):243 [FREE Full text] [doi: 10.3390/vaccines9030243] [Medline: 33799505]
9. COVID Data Tracker. Centers for Disease Control and Prevention. URL: https://covid.cdc.gov/covid-data-tracker/#datatracker-home [accessed 2021-09-02]
10. Griffin JB, Haddix M, Danza P, Fisher R, Koo TH, Traub E, et al. SARS-CoV-2 infections and hospitalizations among persons aged ≥16 years, by vaccination status- Los Angeles County, California, May 1-July 25, 2021. MMWR Morb Mortal Wkly Rep 2021 Aug 27;70(34):1170-1176 [FREE Full text] [doi: 10.15585/mmwr.mm7034e5] [Medline: 34437525]
11. Fisher CE, Boeckh M, Jerome KR, Englund J, Kuypers J. Evaluating addition of self-collected throat swabs to nasal swabs for respiratory virus detection. J Clin Virol 2019 Jun;115:43-46 [FREE Full text] [doi: 10.1016/j.jcv.2019.04.001] [Medline: 30978619]
12. Valentine-Graves M, Hall E, Guest JL, Adam E, Valencia R, Shinn K, et al. At-home self-collection of saliva, oropharyngeal swabs and dried blood spots for SARS-CoV-2 diagnosis and serology: Post-collection acceptability of specimen collection process and patient confidence in specimens. PLoS One 2020;15(8):e0236775 [FREE Full text] [doi: 10.1371/journal.pone.0236775] [Medline: 32756585]
13. Tu Y, Jennings R, Hart B, Cangelosi GA, Wood RC, Wehber K, et al. Swabs collected by patients or health care workers for SARS-CoV-2 testing. N Engl J Med 2020 Jul 30;383(5):494-496 [FREE Full text] [doi: 10.1056/NEJMc2016321] [Medline: 32492294]
14. Wehrhahn MC, Robson J, Brown S, Bursle E, Byrne S, New D, et al. Self-collection: An appropriate alternative during the SARS-CoV-2 pandemic. J Clin Virol 2020 Jul;128:104417 [FREE Full text] [doi: 10.1016/j.jcv.2020.104417] [Medline: 32403007]
15. Altamirano J, Govindarajan P, Blomkalns AL, Kushner LE, Stevens BA, Pinsky BA, et al. Assessment of sensitivity and specificity of patient-collected lower nasal specimens for severe acute respiratory syndrome coronavirus 2 testing. JAMA Netw Open 2020 Jun 01;3(6):e2012005 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.12005] [Medline: 32530469]
16. McCulloch DJ, Kim AE, Wilcox NC, Logue JK, Greninger AL, Englund JA, et al. Comparison of unsupervised home self-collected midnasal swabs with clinician-collected nasopharyngeal swabs for detection of SARS-CoV-2 infection. JAMA Netw Open 2020 Jul 01;3(7):e2016382 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.16382] [Medline: 32697321]
17. Therchilsen JH, von Buchwald C, Koch A, Dam Nielsen S, Rasmussen DB, Thudium RF, et al. Self-collected versus healthcare worker-collected swabs in the diagnosis of severe acute respiratory syndrome coronavirus 2. Diagnostics (Basel) 2020 Sep 09;10(9):678 [FREE Full text] [doi: 10.3390/diagnostics10090678] [Medline: 32916801]
18. Tan SY, Tey HL, Lim ETH, Toh ST, Chan YH, Tan PT, et al. The accuracy of healthcare worker versus self collected (2-in-1) oropharyngeal and bilateral mid-turbinate (OPMT) swabs and saliva samples for SARS-CoV-2. PLoS One 2020;15(12):e0244417 [FREE Full text] [doi: 10.1371/journal.pone.0244417] [Medline: 33326503]
19. Iruzubieta P, Fernández-Lanas T, Rasines L, Cayon L, Álvarez-Cancelo A, Santos-Laso A, et al. Feasibility of large-scale population testing for SARS-CoV-2 detection by self-testing at home. Sci Rep 2021 May 10;11(1):9819 [FREE Full text] [doi: 10.1038/s41598-021-89236-x] [Medline: 33972607]
20. Applying human factors and usability engineering to medical devices: Guidance for industry and Food and Drug Administration staff. US Food and Drug Administration. 2016 Feb. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices [accessed 2021-09-02]

21. Council for International Organizations of Medical Sciences. International ethical guidelines for biomedical research involving human subjects. Bull Med Ethics 2002 Oct(182):17-23. [Medline: 14983848]
22. World Medical Association. World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
23. Evaluation and reporting of age-, race-, and ethnicity-specific data in medical device clinical studies: Guidance for industry and Food and Drug Administration staff. US Food and Drug Administration. 2017 Sep. URL: https://tinyurl.com/bm7uhd3a [accessed 2021-09-02]
24. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. Int J Hum Comput Interact 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
25. Interim US guidance for risk assessment and work restrictions for healthcare personnel with potential exposure to COVID-19. Centers for Disease Control and Prevention. 2020 Oct 21. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/guidance-risk-assesment-hcp.html [accessed 2021-09-02]
26. Lewis JR. Psychometric evaluation of an after-scenario questionnaire for computer usability studies. SIGCHI Bull 1991 Jan 01;23(1):78-81. [doi: 10.1145/122672.122692]
27. QuickFacts: Madison city, Wisconsin. US Census Bureau. 2019. URL: https://www.census.gov/quickfacts/fact/table/madisoncitywisconsin# [accessed 2021-09-01]

Abbreviations

ASQ: After-Scenario Questionnaire
EUA: Emergency Use Authorization
FDA: Food and Drug Administration
IFU: Instructions for Use
N: nucleocapsid
nCoV: novel coronavirus
RNase P: ribonuclease P
RT-PCR: reverse transcription–polymerase chain reaction
WIRB: Western Institutional Review Board

Edited by A Kushniruk; submitted 02.04.21; peer-reviewed by M Popovic, S Lalmuanawma; comments to author 25.08.21; revised version received 13.09.21; accepted 02.10.21; published 14.12.21.

Please cite as:
Strong LE, Middendorf I, Turner M, Edwards V DK, Sama V, Mou J, Adams KC
Usability of an At-Home Anterior Nares SARS-CoV-2 RT-PCR Sample Collection Kit: Human Factors Feasibility Study
JMIR Hum Factors 2021;8(4):e29234
URL: https://humanfactors.jmir.org/2021/4/e29234
doi:10.2196/29234
PMID:34609947

Original Paper

Digitally Disconnected: Qualitative Study of Patient Perspectives on the Digital Divide and Potential Solutions

Maria Alcocer Alkureishi[1], MD; Zi-Yi Choo[2], BS; Ali Rahman[3]; Kimberly Ho[4], BS; Jonah Benning-Shorb[3], BA; Gena Lenti[5], MD; Itzel Velázquez Sánchez[3], BS; Mengqi Zhu[6], MS; Sachin D Shah[1,6], MD; Wei Wei Lee[6], MD, MPH

[1]Department of Pediatrics, University of Chicago, Chicago, IL, United States

[2]Pritzker School of Medicine, University of Chicago, Chicago, IL, United States

[3]University of Chicago, Chicago, IL, United States

[4]New York University Long Island School of Medicine, Mineola, NY, United States

[5]Department of Internal Medicine, University of Washington, Seattle, WA, United States

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

Corresponding Author:
Maria Alcocer Alkureishi, MD
Department of Pediatrics
University of Chicago
5841 S Maryland Avenue, MC 6082 Room C124
Chicago, IL, 60637
United States
Phone: 1 773 834 8927
Fax: 1 773 585 7332
Email: malkureishi@peds.bsd.uchicago.edu

Abstract

Background: As telemedicine utilization increased during the COVID-19 pandemic, divergent usage patterns for video and audio-only telephone visits emerged. Older, low-income, minority, and non-English speaking Medicaid patients are at highest risk of experiencing technology access and digital literacy barriers. This raises concern for disparities in health care access and widening of the "digital divide," the separation of those with technological access and knowledge and those without. While studies demonstrate correlation between racial and socioeconomic demographics and technological access and ability, individual patients' perspectives of the divide and its impacts remain unclear.

Objective: We aimed to interview patients to understand their perspectives on (1) the definition, causes, and impact of the digital divide; (2) whose responsibility it is to address this divide, and (3) potential solutions to mitigate the digital divide.

Methods: Between December 2020 and March 2021, we conducted 54 semistructured telephone interviews with adult patients and parents of pediatric patients who had virtual visits (phone, video, or both) between March and September 2020 at the University of Chicago Medical Center (UCMC) primary care clinics. A grounded theory approach was used to analyze interview data.

Results: Patients were keenly aware of the digital divide and described impacts beyond health care, including employment, education, community and social contexts, and personal economic stability. Patients described that individuals, government, libraries, schools, health care organizations, and even private businesses all shared the responsibility to address the divide. Proposed solutions to address the divide included conducting community technology needs assessments and improving technology access, literacy training, and resource awareness. Recognizing that some individuals will never cross the divide, patients also emphasized continued support of low-tech communication methods and health care delivery to prevent widening of the digital divide. Furthermore, patients viewed technology access and literacy as drivers of the social determinants of health (SDOH), profoundly influencing how SDOH function to worsen or improve health disparities.

Conclusions: Patient perspectives provide valuable insight into the digital divide and can inform solutions to mitigate health and resulting societal inequities. Future work is needed to understand the digital needs of disconnected individuals and communities. As clinical care and delivery continue to integrate telehealth, studies are needed to explore whether having a video or audio-only phone visit results in different patient outcomes and utilization. Advocacy efforts to disseminate public and private resources can also expand device and broadband internet access, improve technology literacy, and increase funding to support both high- and low-tech forms of health care delivery for the disconnected.

(JMIR Hum Factors 2021;8(4):e33364) doi:10.2196/33364

KEYWORDS

telemedicine; digital divide; patient experience; qualitative study

Introduction

Prior to the COVID-19 pandemic, the use of synchronous telemedicine (eg, audio-only telephone or video visits) in the United States was limited and mainly incorporated in specialty fields, such as postoperative care and psychiatry [1,2]. In response to the pandemic, the Centers for Medicare and Medicaid Services (CMS) and private insurers expanded coverage for both video and phone telehealth visits, and telemedicine utilization increased exponentially [3-5] and expanded into primary care [6,7]. While further study is needed to explore the challenges of telemedicine [8], initial studies have found that with certain patient populations and conditions, telemedicine is associated with a number of patient and clinician benefits, including reduced appointment wait times, costs, improved medication adherence and blood pressure control, and high rates of patient and clinician satisfaction [9-19].

However, as telemedicine utilization increased, diverging usage patterns for video and audio-only telephone visits emerged, raising concerns about the widening "digital divide" contributing to disparities in telehealth access [7,20-22]. The "digital divide" refers to a societal division between those who have the technological means to make full use of technology and those who face barriers preventing proper use and benefit [20]. Access to high-speed internet and technology devices (eg, computers, tablets, smartphones) and a degree of digital literacy are required to successfully participate in video visits [7,23]. As health care becomes more reliant on technology-based tools, the digital divide stands to further exacerbate existing health care access disparities.

Studies have shown that patients with lower levels of digital literacy and access to technology are more likely to be from marginalized backgrounds, including older, Black and Hispanic, non-English speaking patients, and those with Medicaid insurance [24]. Further, during the pandemic, patients from disadvantaged socioeconomic backgrounds were less likely to complete video visits and more likely to rely on audio-only telephone visits to access their providers and telehealth utilization data from federally qualified health centers confirmed these findings [7,20,25-28]. Because a third of Medicare telehealth encounters were audio-only phone visits between March and June 2020, significant concerns around worsening health inequalities are raised if reimbursement parity between video and telephone visits is discontinued in the future [4,23,29-31].

While studies have assessed the demographics of the digital divide, none have directly explored patient perspectives on the digital divide and potential impacts. We aimed to interview patients to understand their perspectives on (1) the definition, causes, and impact of the digital divide; (2) whose responsibility it is to address the digital divide; and (3) potential solutions to mitigate the digital divide.

Methods

Setting

The UCMC serves a diverse medically underserved patient population on the South Side of Chicago. During the COVID-19 pandemic, UCMC began offering virtual visits in March 2020 in response to the March 6, 2020, policy changes and regulatory waivers from CMS and provisions of the US Coronavirus Aid, Relief, and Economic Security Act, effective March 27, 2020 [3,7]. From December 2020 to March 2021, we conducted semistructured telephone interviews with adult patients and parents of pediatric patients who had virtual visits (phone, video, or both) at UCMC adult or pediatric primary care clinics from March 2020 to September 2020.

Researcher Characteristics

The research team consisted of 2 faculty physicians (MA and WL), 2 medical students (GL and Z-YC), and 4 undergraduate research assistants (AR, IVS, JB-S, and KH). Interviews were conducted by AR, GL, IVS, JB-S, KH, and Z-YC. Qualitative analysis was performed by AR, JB-S, MA, WL, and Z-YC.

Interview Guide Development

The interview script was part of a larger qualitative interview study focused on understanding patients' overall telehealth experiences during the pandemic. The second half of the interview focused on digital divide perspectives, and the script was developed after a literature review on the digital divide in health care and patient perspectives on telehealth. Results from an internal Press Ganey patient telemedicine survey were used to guide question development. An advisory group of key institutional leadership and stakeholders at UCMC (eg, Vice President and Chief Ambulatory Medical Officer, Associate Chief Medical Information Officer, Advancement Manager for Health Literacy, Diversity and Inclusion) provided feedback on the interview guide and patient recruitment. Patients and family members from the UCMC Ambulatory Patient and Family Advisory Council also provided feedback.

The interview guide (Multimedia Appendix 1) included 4 demographic questions and 6 open-ended questions to elicit perspectives on how patients define the digital divide, its impacts, who they believe is responsible for addressing the divide, and how it could be resolved. A definition of the digital divide was provided to all participants, regardless of their ability to correctly define the concept or not. All research assistants completed pilot interviews, received feedback from the senior authors MA and WL, and revisions were made to the interview script to improve question clarity and focus.

Sampling Strategy

Details on all adult and pediatric primary care patients who had phone, video, or both virtual visit types at UCMC between March 2020 and September 2020 were extracted from the

electronic health record and these patients were eligible for inclusion in the study. Patient demographic data were also obtained (eg, visit type/date, insurance type, primary language, age, sex, race, and phone number).

We purposefully sampled patients to ensure we captured patient experiences for both adult and pediatric patients who had different visit types (phone, video, or both). Patients were randomly chosen until we had representative participants from each subgroup (adult phone, adult video, adult both, pediatric phone, pediatric video, and pediatric both). Participants received up to 2 phone calls to invite them to participate, and oral consent was obtained. We aimed to complete between 25 and 50 total interviews based on prior qualitative studies in our patient population and previous telehealth studies [32-37]. Patient recruitment continued until thematic saturation was reached. All participants received a US $20 gift card to compensate them for their time.

Data Analysis

A total of 54 phone interviews were conducted, digitally recorded, assigned a randomized subject identification number, and submitted for professional transcription. Three research assistants (GL, IVS, and KH) reviewed and deidentified transcripts to ensure accuracy and anonymity. ATLAS.ti 9 was used for qualitative analysis [38]. Using a constant comparative approach, a coding team (AR, MA, JB, WL, and ZC) performed iterative content analysis of 3 transcripts. An additional 9 interviews were reviewed independently and discussed as a group until the code book was finalized and consensus was reached on theme saturation. The remaining 42 interviews were analyzed by AR, JB, and ZC. All coded transcripts were validated by at least one of two reviewers (MA and WL). Both reviewers also independently coded a subset of the interviews, comparing analyses to ensure effective data triangulation.

Institutional Approval

The project conforms with the Standards for Reporting Qualitative Research [39], and was approved as a quality improvement project by the University of Chicago. As such, this initiative was deemed not human subjects research and was not reviewed by the Institutional Review Board.

Results

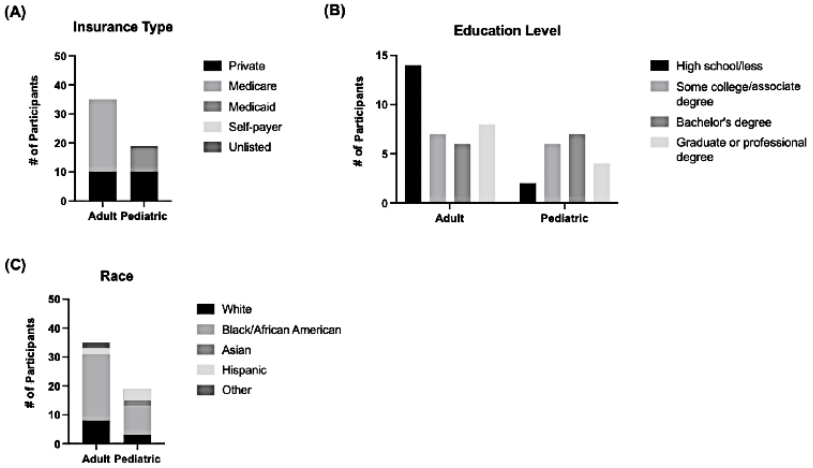
Overview

A total of 216 adult patients were contacted for the study, and 35 consented and completed interviews. In pediatrics, 104 parents were contacted, and 19 consented and completed the interviews. The majority of respondents were adult primary care patients (35/54, 65%), and 35% (19/54) were parents of pediatric patients. Interviews lasted an average of 42 minutes (range 23-79 minutes). Most participants were female (35/54, 65%) and the average age was 55 years (Table 1). The only significant difference between adult and pediatric parent participants was their age (average age: adults: 63.6 years; pediatric parents: 39.1 years; $P<.001$). Patient demographics with respect to insurance, educational attainment level, and race are presented in Figure 1.

Table 1. Respondent demographics.

Demographics	Overall (n=54), n (%)	Adult (n=35), n (%)	Pediatric (n=19), n (%)	*P* value
Race, n (%)				.11
Hispanic	6 (11)	2 (6)	4 (21)	
Asian	2 (4)	0 (0)	2 (11)	
Black or African American	33 (61)	23 (66)	10 (53)	
White	11 (20)	8 (23)	3 (16)	
Multiple/other	2 (4)	2 (6)	0 (0)	
Sex, n (%)				.01
Female	35 (65)	27 (77)	8 (42)	
Male	19 (35)	8 (23)	11 (58)	
Age in years, mean (SD)	54.96 (20.48)	63.60 (19.74)	39.05 (9.30)	<.001
Insurance type, n (%)				<.001
Private	20 (37)	10 (29)	10 (53)	
Medicare	25 (46)	25 (71)	0 (0)	
Medicaid	8 (15)	0 (0)	8 (42)	
Self-payer	0 (0)	0 (0)	0 (0)	
Unlisted	1 (2)	0 (0)	1 (5)	
Highest education level, n (%)				.09
High school/less	16 (30)	14 (40)	2 (11)	
Some college/associate degree	13 (24)	7 (20)	6 (32)	
Bachelor’s degree	13 (24)	6 (17)	7 (37)	
Graduate or professional degree	12 (22)	8 (23)	4 (21)	
Primary language, n (%)				.35
English	53 (98)	35 (100)	18 (95)	
Spanish	1 (2)	0 (0)	1 (5)	
Prior telemedicine visit type(s), n (%)				.03
Phone	17 (31)	12 (34)	5 (26)	
Video	22 (41)	10 (29)	12 (63)	
Phone and video	15 (28)	13 (37)	2 (11)	

Figure 1. Respondent demographics by (A) insurance, (B) highest level of educational attainment and (C) race.



Patient Definition of the Digital Divide

> *I think the digital divide is all the things that COVID has exacerbated. People with money versus people without it. Having access to the internet, makes you able to do all this stuff, and not having access it's harder to do"* [Patient 28]

When asked to define the "digital divide," the majority of patients were initially uncertain of its exact meaning. However, after prompting patients with the definition from our interview script, many recalled personally experiencing the divide and witnessed its effect on family, friends, their communities, and society.

While offering their own definitions of the divide, many patients spoke about the inevitability of technology in every facet of modern life:

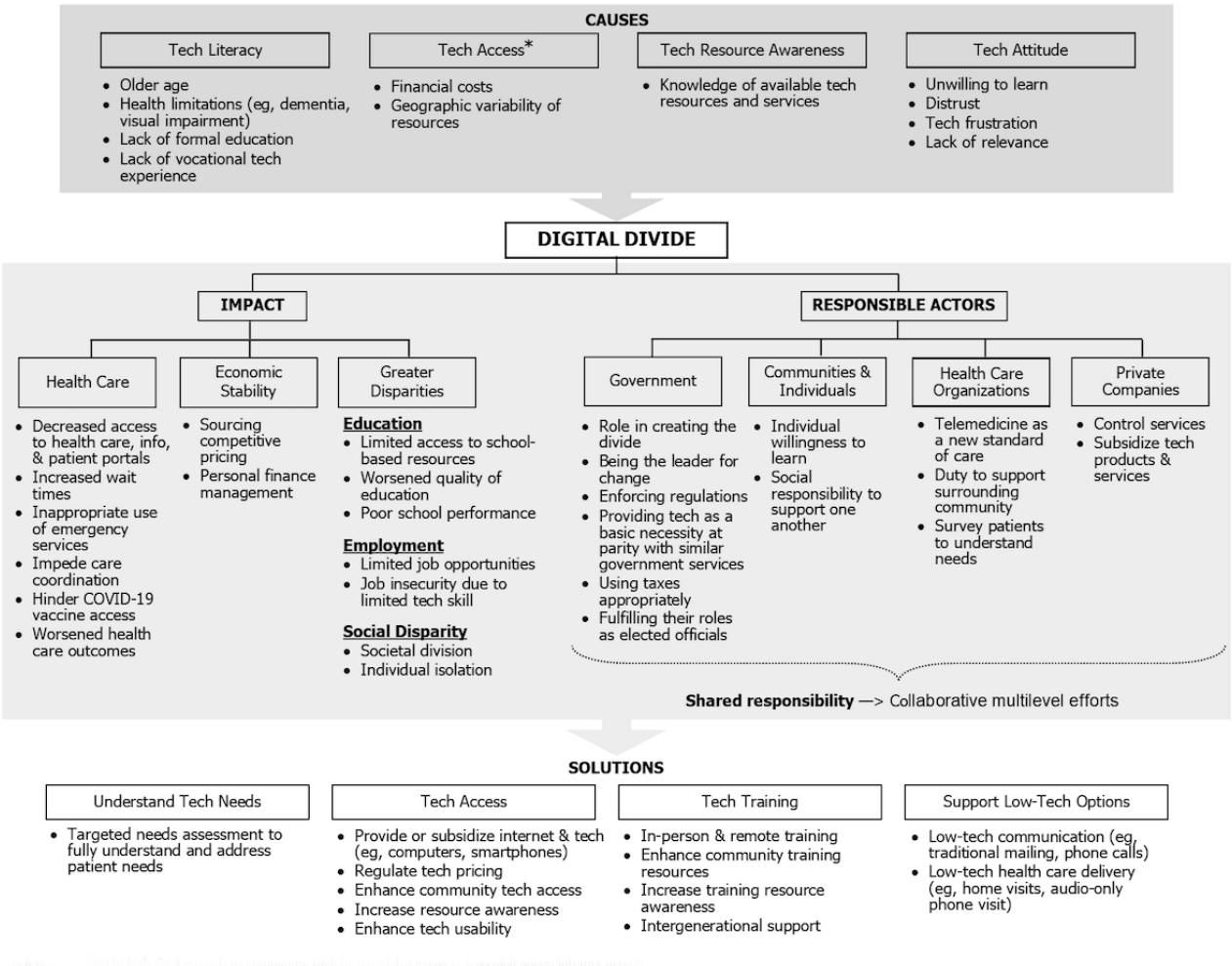
> *Having internet is almost, it's not a right or anything, but it's fundamental in today's society, especially with COVID.* [Patient 28]

Many felt the growing prevalence and necessity of technology in society. Patients recognized how the differential access and ability to use technology created and continues to widen the chasm between those who can and cannot use technology:

> *Well, it's not a good thing because we all know technology is coming faster and faster. To be divided and have people who don't know, it's really not a good thing, because if we are going to be using technology for our medical and our health care, we need to know these things.* [Patient 50]

Having defined the digital divide, 4 overarching themes were identified in the analysis of interview data relating to its (1) causes, (2) potential impacts, (3) responsibility for addressing, and (4) potential solutions. These are explored in more detail below with their respective subthemes and representative quotes in Multimedia Appendices 2-5 and illustrated in Figure 2.

Figure 2. Patient perspectives on the digital divide: causes, impacts, responsible actors, and potential solutions. *Tech access refers to tech devices such as computers, tablets, smartphones as well as high-speed internet access.



Causes of the Digital Divide

Four themes on the causes of the digital divide emerged (Multimedia Appendix 2 and Figure 2): limited technology literacy, limited technology access, unawareness of technology resources, and negative personal attitudes toward technology.

Limited Technology Literacy

Beginning with technology literacy, the most common subtheme cited how advanced age is a major contributor and limitation to their ability to learn and navigate technology. Younger and older patients alike recognized the pervasiveness of technology in many aspects of life, not just in health care. They recognized the challenge for older individuals to keep up with an increasingly technology-dependent society in which computers and smartphones evolve so quickly:

> *At some point, everybody's going to die off who doesn't know how to use them. We got two-year-olds who know how to use it. It's like, eventually, I'm the only anyone who doesn't know how to use it.* [Patient 45]

Cognitive and medical impairments, including memory loss and hearing and visual impairments, were also challenges that contributed to the digital divide among older individuals. One such respondent with dementia explained how her cognitive illness contributed to her isolation from technology and, in turn, where she stood on the digital divide:

> *You know I have dementia...I don't know how to do it. I just don't, and it just bothers me sometimes I just can't. It is very frustrating...I can't put it in words right now but it does make you feel different, with all this technology and not enough personal contact. Well, I can't stop it. I wish it was simple, but it's not going to get better. I can't do what I can't do. If you're not caught up, if you don't know how to do it, sometimes you miss a lot. I think it's unfair. Most people don't want to be a burden to the government or their family either way. They just want to take care of themselves as best they can.* [Patient 36]

Additionally, patients who never had formal training (subtheme 4) or vocational exposure to computers (subtheme 5) found it challenging to stay connected with the rapidly changing nature of technology. These individuals struggled to understand basic technology functions such as sending an SMS text message and were left behind, widening society's digital divide.

Limited Technology Access

The second major cause of the digital divide identified by patients is limited technology access, caused by high financial costs and geographic variation of available services. Patients shared that purchasing technology or an internet plan was a heavy financial burden, particularly for families with multiple children. In other words, cost greatly limits technology access:

> *I mean look at these kids now that are in school. That's why they need to go back to school because you've got kids or multiple children that are supposed to be on the computer in their relative classroom. They have four, five, six kids. They can't share the same computer. They have to be in class at the same time and these parents, who has enough money to get three or four computers around the house? Really, realistically you don't unless you have some nice money. You don't have money for that. No. My sister's a teacher and she's at home trying to be a teacher and then the kids, they have to go to school. It's crazy. [chuckles] It's crazy.* [Patient 6]

Patients also discussed the disparities in high-speed internet access between rural and urban populations:

> *It's socioeconomics for the most part in an urban area...It's being able to afford it and not having enough public resources out there to help bridge that. If you're rural, that's a whole other set of problems that I'm not as familiar with. I know that there are areas that maybe don't have as much cell connectivity, maybe don't have the same internet options. That would exacerbate it. In this area, it's economic. It's being able to afford it.* [Patient 31]

Unaware of Technology Resources

Many patients were unaware of resources such as free or low-cost internet and smartphones, which was identified as a major contributor to the digital divide:

> *Well, part of it is economics, but I think another part of it is a lot of people don't understand what's available to them. Because they have programs to cover, but a lot of people didn't know how to take advantage of it. They were just out there on their own with nothing.* [Patient 17]

Negative Personal Attitudes Toward Technology

Many patients stated that older individuals are generally reluctant to use technology. Older patients were often frustrated with learning new technology, had feelings of distrust, and were hesitant to integrate it into their lives particularly when the majority of their daily functions did not require technology use. However, even though less technology-savvy older individuals avoided technology use, they did recognize the way of the world was now increasingly technologic:

> *I think when you didn't get in on that ground floor, some years ago when they started coming and they started coming so fast, it put a little fear into them, and they're like, 'I don't want to be bothered with that.' It just took off in a whirlwind and went faster than they were going. I'm guilty too. Probably when that was going on, if I had been saying to my mother, 'Oh, come on let me show you, let me show you.' She might be a little more open to it. I have a friend, a very good friend, her dad is 98. He lives on his own, his mind is clear, doesn't have a lot of illnesses. It's like, 'I don't want to be bothered with that.' Now they don't have the confidence. It's like, 'Oh, I'm too old for that.' That kind of thing. I'm too old for it too.* [Patient 50]

As a result, patients believed that the digitally disconnected and less knowledgeable individuals were at a disadvantage, more isolated, and left behind.

> *Society has not accepted the world of non-technological people. I think that they are missing out and I'm really big on what's going on with them, with the seniors. They don't have people to do it for you. My dad has his children, my aunt doesn't. I have*

> *done a lot of things for everybody up until the digital thing prevented me, it hold me down.* [Patient 19]

Impacts of the Digital Divide

Patients described the digital divide impacting (1) health care, (2) economic stability, (3) education, (4) employment, and (5) social disparity (Multimedia Appendix 3 and Figure 2).

Health Care

Eight subthemes were identified related to the impact of the digital divide on health care (Figure 2). Most were related to the expansion of technology-dependent virtual care during the pandemic and how low technology resources and literacy hindered the ability to engage in necessary health care activities such as video visits. One patient stated:

> *That would mean that there's a lot of individuals who are not receiving adequate medical care right now during the pandemic because they don't have access to those resources.* [Patient 10]

Some patients felt the digital divide also increases visit wait times. Specifically, they believed that those who are able to conduct virtual visits have quicker access to care including responses to their patient portal messages and increased availability to virtual visits. Conversely, those who are unable to access these online resources had to wait for return phone calls and in-person visits, which are not as readily available.

There were concerns regarding inappropriate overuse of emergency services. Without a virtual option, patients will be more likely to visit the emergency department, which increases the hospital's burden and possibly inappropriate care:

> *It would probably cause maybe overpopulated emergency departments because now instead of being able to have video visits. I probably would think that it would lead to people not really getting the health care that they need.* [Patient 35]

The digital divide can also limit access to online patient portals. Without access to these tools, less technologically able individuals experience challenges in care coordination such as scheduling visits, communicating with their clinicians, and facilitating referrals and tests. Patients also remarked that these portals are not just useful for themselves but also for children's and elders' parents and caretakers. Furthermore, lack of internet access impedes general knowledge seeking and access to online, reliable, and high-quality health information.

Patients also found that the divide can worsen personal health care outcomes because of limited opportunities and resources to coordinate health care needs. Lastly, patients recalled how limited low-tech outreach efforts for the COVID-19 vaccine and online scheduling portals posed significant challenges for less technology-savvy individuals:

> *They (seniors) can't handle that, it's too much. I think that senior people should have a little special consideration seeing that because of the fact that a lot of them are not technically technology-oriented. Why you don't call them, and tell them when they can get an appointment for the vaccine? Why do they have to go online and look for an email, when they don't do computer? See, I have somebody facing this right now.* [Patient 19]

Economic Stability

One subtheme focused on how the internet allows individuals to quickly source and compare prices of basic goods and services. The second subtheme was the impact of the digital divide on the management of personal finances and accounts, which are predominantly done online today. For example, some patients noted that during the pandemic, managing aspects of their basic utilities were only accessible online, which created a frustrating and unfair experience for digitally disconnected individuals and families.

Education

Impacts of the digital divide on children's educational experience were especially prominent due to remote learning during the pandemic. During the pandemic, the lack of access to on-site school–based computers and internet limited children's ability to gather resources needed for their learning and school work:

> *We had kids who don't have computers, whose moms were sitting outside the school just so they could have WiFi so their kids could learn.* [Patient 45]

Society's increasing dependence on technology also increases financial burden on families with several children. Larger families with multiple children need to spend more to provide basic educational supplies, which are now typically expensive technology items and services. Patients discussed how the divide impacts the quality of education in socially and economically disadvantaged areas, contributing to poor academic achievement. Several patients spoke about how children in communities of color, particularly those on the South Side of Chicago, are more technologically challenged than other children from privileged areas due to less exposure to computers. Thus, the transition to an entirely technology-dependent remote learning environment during the pandemic further exacerbated these disparities and resulted in poorer academic performance:

> *Well, the (South Side) kids, they have a harder time with just being on the computer. I'm sure some of them have figured it out, but on a day-to-day basis, even in school, they're not on computers like some kids in the suburbs and kids who are more privileged. They're far ahead of some of our kids. This has put our kids behind. That's what I feel.* [Patient 47]

Employment

The inability to access the internet also limits employment opportunities such as limiting one's search for jobs and professional networking. This reinforces a societal divide where the less educated and technologically able are limited to lower paying, more manual jobs:

> *It's a privilege to be able to have technology and have that resource readily available. Those who are able to have access to that, it's in a sense like a sense of superiority. It's just a complex of it's a hierarchy in a sense where more people are more educated and*

> *those who are less educated. When you're more educated, you're able to be more involved or get higher-paying jobs. In that manner, it creates a divide and a conflict.* [Patient 53]

Furthermore, the divide threatens job security. Patients noted employers prefer to retain more technology-savvy employees, and computers and machines increasingly replace manual workers.

Social Disparity

Lastly, patients highlighted how technology contributes to societal disparity by promoting division and isolation of those that are less technology savvy. More recently, the pandemic and an increasingly technology-reliant society have further isolated and ostracized digitally disconnected individuals:

> *I think (the digital divide) it's all the things that COVID has exacerbated. People with money are able to get access versus people without it. It's basically having access to the internet, makes you able to do all this stuff, and not having access it's harder to do, I think. I think we're seeing what that means, but I think it does exacerbate conditions like whatever the current state is and just speed things up.* [Patient 28]

Patients also highlighted the duality of technology: technology can positively and powerfully empower the literate while repressing and leaving the less tech savvy with a more difficult life:

> *That would set a lot of people apart. It really divides people because you've got to have access to a computer now in order to do the smallest things. You really do. You've got to log onto this to do whatever the application-- I haven't been to the library in a long time, but I bet you have to have a computer now in order to do some things at the library. What's it's going to do I think it's really going to divide people pretty soon if technology can change the world, it's just me saying this, but I think that if technology can change the world the way it is...I think it'll really put a rift between, instead of the high class and the low class, it would be a divide between the illiterate and the literate. Computer literates and illiterates and literates. Well, I think it means that you're dividing to go into two classes. One that is more privileged, probably have an easier life than the other.* [Patient 37]

Responsibility for the Digital Divide

Patients had clear perspectives on whose responsibility it is to address the digital divide, and these were organized according to 6 main actors and their explanatory subthemes (Multimedia Appendix 4 and Figure 2).

United States Government

By far, the US government was thought to be primarily responsible for addressing the digital divide. Patients believed the government played a role in causing the digital divide, thus they should address it. Additionally, patients told us that only the government has the power to legislate and enforce regulation to protect individuals from unfair business practices:

> *I think that has to do with the government. I think that's a government issue because I think they the ones that could actually make them (companies)-- they could actually put a cap on all of this.* [Patient 2]

Several patients emphasized a subtheme related to the need for the federal government to stand out as the leader for change and initiate a top–down plan of action between local governments, businesses, and community organizations to address the divide:

> *Of course, we know our government needs to be a part of that. That's from The White House, all the way down to our local government. We need those people to be involved because to bring these resources to the forefront, we certainly need money.* [Patient 50]

With the rising prevalence of technology, patients viewed technology and internet connectivity as a basic life necessity like food, transportation, and clean water. Given that the government addresses issues such as food insecurity, they should be similarly responsible for addressing digital insecurity and providing equitable technology access to citizens:

> *I guess, yes, if that were a project like the highway system or something, or the interstate system, or health care. It's so central to life. It affects your quality of life if you don't have access to it. It's like the water. It wouldn't be right if people have limited access to food or water. People should have access too.* [Patient 28]

Patients also believed that elected individuals and taxes should be used to support individuals, communities, and social programs with technology access and training:

> *We should work to make internet available to everybody. I know the federal tax is designed to do that. We pay extra on our phone bills to pay for internet for people that don't have it...I am definitely of the type that I would rather pay a bit more in taxes and see absolutely everyone have their needs met. I think it's on the government. If we pay taxes, it's for services and being able to function and be in school as you are required to. The government should make that available.* [Patient 5]

Individual Responsibility

Patients recognized that aside from individuals with cognitive impairments, people have a choice in learning how to use technology and seeking access to keep up with the digital world. There needs to be a component of individual willingness to bridge the divide:

> *It'll still be up to that person if they want to make that change and catch up with the world. It's up to them whether they want to learn. If they don't want to learn, you can't make them. It would be their responsibility. It's just like saying the video visits versus going there in person. If they don't want to learn technology, all*

right, then you got to go down there in person. [Patient 49]

Health Care Organizations

Health care organizations were cited as a responsible actor for the divide, especially because technology-based care such as video visits are quickly becoming commonplace and a part of the standard of care. Patients believed that health care organizations have a duty to support their local communities with not only medical services but also technology resources and training because they are in the unique position to directly survey their patients and understand their technology barriers:

> *Because they're your health care provider. They can't really fix the digital divide that's not related to health care, not the University of Chicago. But in relation to health care, then I think they should be responsible for that. Because I know that a lot of the seniors are not getting the same, because of the fact that they are not technological.* [Patient 19]

Private Companies

Patients expressed that private companies should address the digital divide because they control technology services and resources and have the potential and responsibility to allocate them fairly. Several patients also recalled how private companies denied or did not offer subsidized services to individuals who needed financial assistance. Specifically, one patient stated that part of their own internet bill payments could and should be used to help support services for others who cannot afford internet plans:

> *Maybe individuals can put pressure on the companies, or we can in our bills agree to pay more because- to help reduce the cost to make sure so that people provide it but then have people sharing, which I think is the next best thing.* [Patient 46]

Communities

Patients expressed the need for social responsibility and technologically able family and friends to support individuals who struggle to access and use computers and smartphones:

> *I think that the community where we live at. There are programs out here, it's resources that assist with computer classes that get people used to using a computer or even just to know that normal functions of a computer, community resources to assist with that.* [Patient 20]

Shared Responsibility

Lastly, there was a profound recognition that the divide is a real and serious threat to the well-being of our communities and nation. Given the scope and significance of the divide, everyone, including the government, organizations, companies, communities, churches, libraries, and individuals, plays an important role in addressing it:

> *The digital divide will have to be fixed by everybody. Just like anything else; in the US, when it's time for the big push, everybody has to cooperate. Not just one or two people, or one or two agencies. Government can't do everything. If they could, honey, we'd be a Socialist Party for real and this would be France, but they can't do everything. It's going to take everybody. Every church, every institution, every computer company. The digital divide is not a joke. It is for real just like food deserts; no grocery stores in certain areas for miles, and what if you're on a bus? If you can at least get a phone, then most people can cross that digital divide.* [Patient 4]

Potential Solutions to the Digital Divide

Four themes were identified as possible ways to overcome the digital divide (Multimedia Appendix 5 and Figure 2).

Understanding Technology Needs

The first and most critical step in solving the divide is understanding it. Patients believed that an initial targeted technology needs assessment of communities and health care organization members would facilitate a baseline understanding of what their unique needs are:

> *One-on-one surveys with people in various ages and ethnic backgrounds to see how they feel and what their needs are. In order to close the gap, you got to see what you need. You've got to put the information out there and find out what people really need. How many 70-year-olds, or 90, or 40-year-olds need these resources? When you don't know, how can you fix it? You've got to know what a person's needs are before you could fix it. I think starting there, we would get a lot of answers and do things differently.* [Patient 50]

Ensuring Access to Technology

Patients proposed directly providing devices such as computers and smartphones, as well as internet connectivity, to all. Again citing technology access as a basic necessity particularly during the pandemic and for virtual health care, patients called on a variety of actors (eg, government, health care organizations, private organizations, social infrastructure such as schools) to step in and provide the necessary tools to individuals. With many patients unable to afford the required technology to participate in today's digital world, universally providing the technology is one way to ensure equitable access to everyone:

> *They have to be able to provide everybody that sort of resource. Just if everybody has equal access to opportunities also that are provided through the internet. Free access to Wi-Fi and providing something to access that, like a computer or a tablet. Providing the resource for these people to be able to access. Even libraries are not enough or some small things like that are difficult.* [Patient 53]

However, recognizing this may not be feasible. Many respondents proposed providing subsidized technology resources to regulate technology costs, thereby improving access to those who need them most:

> *Some people can't afford the internet. Some people is just living off of once a month check or some people not getting any income at all, so how would they go*

> *about paying their internet bill? To me, for people like that, and people that can't afford it. I think they should have a program for them, where they should be able to get it for free because they know it's a need that they need.* [Patient 25]

Enhanced shared technology resources in new or existing community settings such as providing free shared use computers or phones in doctors' offices, public areas, internet cafes, and community centers can help increase technology access:

> *If there is someplace that we could just use their computers that are already up and specialized...I would go in a heartbeat. You just need help and when you need help, you do whatever you have to do...Give us an option of being able to come into a room or some area where you could come in and use their actual equipment which might be better for all of us.* [Patient 6]

Many patients specifically commented on how libraries can help bridge the digital divide because they are widespread, play an active role in underserved communities, and offer open access to technology devices and internet connectivity. However, due to COVID-19, patients expressed frustration in decreased library availability and called for the need to re-evaluate and expand library services, such as revising opening hours, providing socially distanced areas to use technology devices, and providing patient care pods where virtual visits could be privately conducted:

> *Everything has changed. They need to be aware. The library is open from 10:00 to 6:00 or 10:00 to 5:00. Like I said, the library used to be open up to nine o'clock at night, but I know it's not anymore. I need my Zoom (video visit) call to be at 11 o'clock. Why? Because the library doesn't open up until 11 o'clock. I know McDonald's is open, but McDonald's you can't sit in McDonald's now. There's no place that you can go and actually sit but the library. You can only be in the library a maximum of an hour, I think. You know what I'm saying? It's a lot of work but--It's important to keep people alive because I think that's what our main focus should be right by now is to keep people alive, healthy and safe.* [Patient 48]

To ensure access to technology, there needs to be an enhanced awareness of one's available resources. Lack of technology resource awareness can be just as prohibitive a barrier as not having the device itself. Again, with closures and accessibility restrictions due to COVID-19, overcoming the digital divide requires active communication of available technology resources to connect them with the individuals who need them most.

And lastly, patients expressed that improving the patient-facing functionality and usability of technology in health care would facilitate technology usage. Suggestions such as more patient-friendly online portals, the ability to share visits more easily, granting portal access to more than just 1 individual (eg, to 2 parents), and video visit platforms that were easier to navigate were all cited as potential solutions to bridging the technology divide. Improved user interfaces could improve how patients use these tools in their own care or to assist family and friends with their care.

Technology Training

The second overall solution theme to the digital divide was the provision of technology training to improve digital literacy. One subtheme called for in-person training because it is more relatable and easier to understand. Patients emphasized that the instruction needs to be "simple, simple, simple. It got to be simple (Patient 36)" particularly for older adults or individuals with cognitive impairments such as memory loss. Patients recommended educational institutions such as universities and health care organizations as good venues for hosting workshops, ongoing classes, and even a dedicated technology help desk in clinics where patients and family members could learn how to navigate their online patient portals in-person, conduct a video visit, and use technology in general:

> *They have classes for everything at the university. Just like they set that up, set a class up. When people come in the hospital, they can go, "Oh you know what? Oh, they having technology class. Since I'm here at the clinic that day, oh let me see they had a class that day?* [Patient 32]

Additionally, synchronous (eg, a phone line) and asynchronous (eg, preparatory instructional videos and written information) remote learning resources can help patients overcome technology issues related to video visits or the use of patient portals. To troubleshoot potential challenges in advance, some also suggested virtual practice sessions and opportunities to access their video visit platforms prior to their actual appointment. Of all the proposed remote assistance recommendations, the most common was the establishment of a dedicated technology helpline and help desk for patients, particularly because it can be difficult to reach someone to resolve technology issues via the main clinic line:

> *It'd be helpful if the doctor probably had a tech department...Where you don't have to wait for your doctor to call you back or your nurse to call you back in regard to getting on to your appointment.* [Patient 54]

Intergenerational help from technology-savvy family members such as children and grandchildren was another commonly cited way to overcome issues with technology literacy. Many patients, particularly older adults, reported having a greater dependency on tech-savvy younger relatives and friends, highlighting the importance of intergenerational assistance to help less technology literate patients navigate technology:

> *I can't do the latest model phone they have out. My granddaughter won't give me one of those but my great-grandchildren have them. That's who teach me how to work on the computer and it's stuff my great-grandchildren, the little ones. If I have trouble with my video visit, I have two "greats" sitting right here.* [Patient 8]

Similar to the subthemes about enhancing access to technology and increasing awareness of those resources in the community, most patients also emphasized expanding technology training

and resource awareness to address digital literacy gaps. Extending this, several patients also envisioned having technology champions and coaches directly in their community. These individuals could then volunteer to use their knowledge of computers and smartphones to educate others in their communities:

> *It actually has two advantages to it. One, it is creating a job for two people to come in to teach, definitely be the patient that they will benefit from it. It does get another people that will be able to talk and communicate with other people, so they can teach each other because that's how we learn; we learn from one another. If they have a technology class, and maybe they're saying, okay, this is what you need to do to take classes. Maybe they don't know how to set it, maybe I do. Okay, you know what? I'll show you how. The teacher is learning from the students and the students are learning from the teacher too. Being knowledgeable, it balances out and nearly everybody feels needed.* [Patient 2]

Supporting Low-Tech Health Care Modalities

The final theme rested on the recognition that patients who could not or would not cross the digital divide will continue to suffer in terms of their health care access. For these individuals, the need to provide continued equitable access to quality care is paramount. This quality care includes the continued use of low-tech communication modalities such as phone calls and postal mailing to convey important information such as how to get their COVID-19 vaccine.

> *I think that senior people should have a little special consideration seeing that because of the fact that a lot of them are not technically technology-oriented. Why you don't call them, and tell them when they can get an appointment for the vaccine? Why do they have to go online and look for an email, when they don't do computer? See, I have somebody facing this right now. I think that they should reach out to those people. They shouldn't have to reach out to them. Then some of them are not in line for it because of the lack of technology and I think that's wrong.* [Patient 19]

Audio-only telephone visits are particularly important for individuals with physical or cognitive limitations, given they described completing a video or in-clinic visit challenging if not impossible.

> *Some of it is because of COVID. Sometimes it's because the cases it seems, I'm not feeling good and depending on what's going on with us, it can become challenging to get him up and get him dressed and get him to the doctor. It's (audio-only telephone visits) talking to your doctors, it's not a joy call. You know what I'm saying? This is not where you stand up to go out to dinner, we're talking about life and death here.* [Patient 2]

Furthermore, low-tech health care delivery in the form of home visits is an important adjunct to care for select patients, and provided a critical health care lifeline that should be supported beyond the pandemic:

> *My dad and I live together and he's elderly, he struggles with traveling. My thing would be something that would be available for those who are disabled, obviously the elderly who are not in a nursing home or hospice. You know what I'm saying, nursing home care or independent care where professionals can come in and provide the vaccine. They can set up appointments. Essentially, it would be like DoorDash but for vaccines.* [Patient 48]

For many, the pandemic exacerbated the isolation of older and technology-challenged individuals. One patient commented on the need to not only continue providing these low-tech forms of health care and communication but health care organizations also need to increase their efforts and proactively reach out to our society's most vulnerable individuals:

> *This is technology going on, a lot of them are lost. They have no idea because they never worked on the computers. They never had to. Our seniors get lost in the system because they're senior and they're older and they don't have anybody to come and help them. I know somebody that ain't seen their doctor in two years but the doctor never reached out to them either and I was like, well, that's not a good thing. You're the doctor and your patient is a cardiac patient so if you haven't seen your patient in three months wouldn't you have your nurse call and say, 'You know what we ain't see miss so an so because she has respiratory problems cardiac problems. We need to reach out to her. Because maybe there's something going on. If she doesn't have anybody to come see about it and she can't remember to call the doctor because she's feeling bad, then what's going to happen? They'll find her dead in her apartment and then they'll go, 'Well, nobody called, nobody checked.' I know this is a lot of work but it's important to keep people alive because I think that's what our main focus should be right by now is to keep people alive, healthy and safe.* [Patient 2]

Discussion

Principal Findings

This is the first study to directly explore patients' perspectives on the digital divide and capture them qualitatively in their own words. While an individual's technologic ability and access were noted as fundamental causes of the digital divide, other factors such as lack of awareness of community resources often disconnected support from vulnerable individuals. Furthermore, while advanced age has been cited as a contributing factor to the divide in prior demographic studies, patients also noted that health limitations such as cognitive decline and memory loss increase technologic isolation [7,20-28]. The resultant lack of familiarity and understanding along with rapid advances of technology leave many older individuals and those with cognitive impairments feeling frustrated, distrustful, and unwilling to learn. The staggering prevalence of dementia in

our society has left many home-bound and reliant on family and friends for care [40-43]. Given estimates that the number of individuals living with dementia will rise from 55 million in 2019 to 78 million by 2030 and 139 million by 2050, these are important causative factors to account for as our population's median age and cognitive illnesses continue to increase [44].

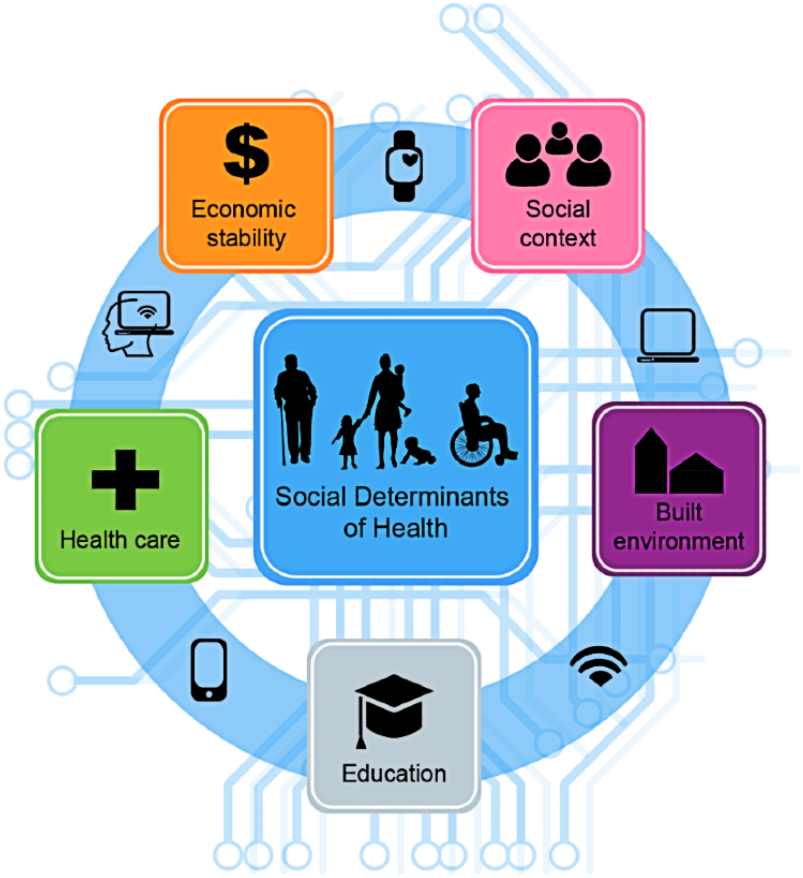
Patients also thought broadly about the divide, recognizing impacts beyond health care (eg, employment, access to nonmedical information, and day-to-day functions critical to individual economic stability such as personal financial management). Education in particular was at the forefront of patient's minds, primarily because of the advent of remote schooling. However, it is important to know there are many studies that have shown a positive correlation between education and quality and longevity of life [45]. Some patients spoke about the "Homework Gap," referring to the lack of equitable student access to high-capacity broadband at home [46,47]. While 93% of people in US households with school-age children reported their child engaged in some form of "distance learning," results from the ongoing US Census Bureau Household Pulse Survey showed that children in high-income households used online resources at higher rates than those in lower-income households [48]. When the pandemic began, 15-16 million K-12 students did not have adequate access to the internet, and by January 2021, up to 12 million students remained under-connected [49]. Even when students did have internet, lack of service provider competition in underserved areas caused unfair pricing and digital redlining, which is the systematic exclusion of low-income neighborhoods from fast broadband service [50,51]. Additionally, an estimated 75% of state and local efforts to bridge digital divide enacted during the pandemic to connect students are due to expire in the next 1-3 years [52]. Based on our patients' experiences and the direct relationship between educational attainment and wellness, it is critical to ensure temporarily connected but underserved students are not left digitally unsupported again [45].

While technology can alleviate disparities among vulnerable populations, it can also create barriers that perpetuate inequalities [20,53-58]. Patients recognized this dichotomy, and because of our society's dependence on technology, patients thought the far reaching effects of the digital divide not only worsened pre-existing societal disparities but also excluded and isolated less-resourced communities. As a result, responses illustrated that technology access and literacy were not only viewed as a basic 21st century right, but patient experiences showed it to be also inextricably intertwined with all 5 social determinants of health (SDOH): (1) health care access and quality, (2) neighborhoods and built environment infrastructure, (3) social and community context, (4) economic stability, and (5) education access and quality. Based on our patients' experiences, it is clear that technology is not the sixth domain on the list of SDOH [59-61]. Rather it is a major controller of every SDOH and the environment in which they can be accessed fully (Figure 3). For example, a stable internet condition and computer are prerequisites for many forms of education. Without these, individuals are unable to take advantage of all that their education offers and often have to undertake cumbersome efforts or greater cost to keep stride with their technology-literate peers. Insufficient technology access and literacy are the drivers of the various SDOH, profoundly influencing whether they are functional or dysfunctional and impacting one's overall health and quality of life. Examples such as this reflect the modern reality that patients described—one in which technology is a gateway to health and wellness, and thus equivalent to other basic needs—and was the basis for their call for solutions to support technology use.

Patients had insightful solutions to begin supporting technology as a basic need, beginning with obtaining a better understanding of patients' technical needs. Conducting a needs assessment and screening patients for their technology needs were important in the health care setting, which fits well within the role and function of the medical home [62-64]. Health care organizations are in an ideal position to develop strategies for overcoming barriers to technology in medical care. However, this is not possible until greater knowledge about patients' technology literacy and access is determined.

A number of creative ways were proposed to support technology access, ranging from direct provision and subsidization of devices to regulated pricing. Patients described the phenomenon of "Parking-Lot Wi-Fi," which is people sitting in parking lots of shuttered libraries, shops, and schools to connect to their only source of Wi-Fi internet access [65]. Squatting in proximity of public Wi-Fi has become so commonplace that states have begun publishing parking lot maps for residents without home internet access [66,67]. However, even in areas where federal internet service maps indicate broadband access, there are still pockets within these areas that lack affordable, high-speed internet [68]. These communities, known as internet or digital deserts, are often in urban areas, and a Census Bureau Survey showed that 3 times as many households in urban areas remain unconnected as in rural areas [69,70]. Although there are high-cost reforms in places such as the Connect America Fund, which provides funding to service providers that commit to offer voice and broadband services to fixed locations in unserved high-cost areas, digital deserts and redlining continue to exist [71].

Figure 3. A conceptual model of patient perceptions of technology access and literacy, as it relates to the various social determinants of health.



Ensuring technology access addresses only 1 facet of the divide, patients recognized improved technology design and training are also critical for improving technical literacy. As an example, patients asked for improved functionality of existing health care technology such as patient portals and video visit platforms. Calls to incorporate the patient perspective in technology design to support not only clinicians but also patients is not a new concept [42]. However, patients are currently asking health care organizations to establish a separate set of resources such as a patient-facing addition of the information technology department to directly support patients with telehealth use. In the face of the digital divide and increasing dependence on telehealth, participatory design and support are more important than ever to consider. Furthermore, given our aging population and the cognitive challenges they often face, providing a simpler and less overwhelming version of these platforms is needed. For example, interfaces with features such as voice activation, memory aids to improve ability to remember important tasks, and easier access sharing with caregivers can make navigation easier and promote a higher level of technical independence [42,72].

Libraries were often cited as a key potential solution to the divide. Libraries hold great promise in helping solve the overall literacy challenges faced by many US adults [73], as well as in addressing technology access and literacy barriers [74-79]. Given their existing social infrastructure and location in communities of need, libraries are well positioned to become part of a more equitable ecosystem of learning, providing access to knowledge, resources, and training that may not otherwise be accessible to people with lower incomes [78,79]. Unfortunately, the pandemic has greatly limited access to many of these spaces and supports, leading patients to ask for a continued re-examination of the role and services that libraries provide. They and other social infrastructures such as churches, senior centers, schools, and parks are important shared spaces in our communities [75]. These spaces have the potential to not only pivot and repurpose existing services but also innovate and collaborate in new ways to bridge the divide and better meet the evolving needs of their communities. To facilitate this transformation, policymakers must invest in efforts such as the E-Rate Program, which was started in 1996 to help public schools and libraries cover the cost of internet access in their buildings. Programs such as the E-Rate can enable libraries and other social infrastructures to take stock of how they are used within marginalized communities and invest in efforts to bring technology access and training to underserved households [76-79].

Beyond providing enhanced technology resources, patients also recognized it was equally important to identify, organize, and connect these services and resources to disconnected individuals. Efforts such as NowPow, a personalized community referral platform that draws on a comprehensively sourced and updated community resource directory, allow clinicians to connect patients to needed medical or social self-care resources [80,81]. Comprehensive, technology-based community asset census mapping will increase awareness of these services and combat the decentralization of community resources by allowing

organizations to work synergistically toward their shared goals of individual empowerment and wellness.

Lastly, patients called for the need to support low-tech health care solutions (eg, audio-only phone visits, mailing information, and home visits) in the wake of the digital divide. This was made especially clear by patients who had difficulty navigating video visit platforms and COVID-19 vaccination efforts through patient portals and online scheduling [82]. While some individuals were open to attempting these new and unfamiliar forms of health care, especially if intergenerational support was available, many were not willing or able to. In these situations, it was necessary to support their continued access to care as a basic right for health care equity.

Supporting low-tech, audio-only virtual care has become an increasingly relevant concern, especially if reimbursement parity between telephone and video visits is discontinued and audio-only phone visits are no longer reimbursed [7,23,30,31,83-86]. Furthermore, if audio-only visits are not supported, those unable to navigate a video visit may defer care altogether—a common occurrence among patients during the pandemic [87,88]. Telephone visits support high-quality care, particularly in primary care and community health settings. Since 2010, the Veterans Health Administration has incorporated scheduled telephone visits into their patient-centered medical home model to improve care access and efficiency [84,89]. In studies with seniors and in mental health settings, audio-only phone visits were as effective as video in resolving urgent and nonemergent needs [90,91]. And in safety-net populations, telephone visits during the pandemic increased access to care, reduced wait times, and in certain circumstances, offered high quality of care comparable to that of video visits [28]. While additional study is needed to compare experiences and outcomes of video and audio-only visits, the potential benefits of video visits will not be realized if patients cannot navigate the technology. As such, continued support of low-tech but high-value care and communication is needed. Further, innovative approaches are necessary to overcome some of the challenges of conducting research with socially disadvantaged and technologically isolated groups, to increase their voice and representation in health and medical research [92].

Limitations

Our study has several limitations which are important to note. While we conducted our study with both adult and pediatric parents, this is a single-institution study and we only included patients that had prior virtual visit experience (phone or video visits), both of which may limit generalizability to other patient populations. It is important to note, however, that surveying this population gave us access to patients who are on both extremes of the digital divide: those that may not have had the technology access and literacy to conduct video visits and had audio-only phone visits, as well as those who were more digitally literate and connected and able to have video visits. Another limitation to generalizability is the large percentage of our study population with higher education, therefore representing a more socioeconomically wealthy group. However, it is important to note that nearly one-third of our respondent population had an educational attainment level of high school or less, and that the majority of our respondents had Medicaid or Medicare insurance. Furthermore, we only solicited the views of primary care clinic patients; results may differ when interviewing specialty clinic patients. Additionally, our low response rate may have contributed to a nonresponse bias. Social desirability and recall bias may have impacted patient responses to our interview questions.

Conclusions and Future Directions

Patients are keenly aware of the digital divide and how it disparately impacts health, work, education, community and social contexts, and personal economic stability. As such, digital access and literacy are not merely another SDOH, they are in fact drivers of each SDOH and are fundamental to their function or dysfunction (Figure 3). Given the complexity of the divide, a shared responsibility between the government, private sector, social infrastructure such as health care organizations, community organizations, public services, and individual citizens is needed. Community and organizational needs assessments are essential to identify and target the most impactful interventions. Providing technology is a necessary first step; however, solutions must reach beyond access alone. Development and dissemination of technology literacy training programs and increasing awareness and coordination of available resources will take time. Importantly, even with all of these efforts, some will never cross the divide. For these individuals, it is necessary to continue to support low-tech means of communication and health care delivery to prevent further isolation and widening of the digital divide.

Future work is needed to understand the technology needs of digitally disconnected and excluded individuals and communities. As telehealth continues to be integrated into clinical care delivery models, studies should explore how video versus audio-only phone visits impact patient experiences and outcomes, as well as the role of digital navigation to proactively identify and address technology needs. Advocacy efforts should focus on the utilization of public and private resources to address issues such as expanding device and broadband internet access, improving general and health technology literacy, and advocating for policy to continue supporting both high- and low-tech forms of health care delivery for those unable to cross the divide.

Acknowledgments

This study was funded by the Comer Children's Hospital Development Board, as well as the Bucksbaum Institute for Clinical Excellence and the Center for Healthcare Delivery Science & Innovation at the University of Chicago Medicine. We thank our patients for taking the time to discuss their perspectives with us, as well as our graphic artist Zi-Ning Choo for her assistance with the conceptualization and design of Figure 3.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Digital divide interview script.
[DOCX File , 19 KB - humanfactors_v8i4e33364_app1.docx]

Multimedia Appendix 2
Causes of the digital divide.
[DOCX File , 16 KB - humanfactors_v8i4e33364_app2.docx]

Multimedia Appendix 3
Impacts of the digital divide.
[DOCX File , 18 KB - humanfactors_v8i4e33364_app3.docx]

Multimedia Appendix 4
Responsibility for the digital divide.
[DOCX File , 16 KB - humanfactors_v8i4e33364_app4.docx]

Multimedia Appendix 5
Potential solutions to the digital divide.
[DOCX File , 20 KB - humanfactors_v8i4e33364_app5.docx]

References

1. Bashshur RL, Shannon GW, Bashshur N, Yellowlees PM. The Empirical Evidence for Telemedicine Interventions in Mental Disorders. Telemed J E Health 2016 Mar;22(2):87-113 [FREE Full text] [doi: 10.1089/tmj.2015.0206] [Medline: 26624248]
2. Chao GF, Li KY, Zhu Z, McCullough J, Thompson M, Claflin J, et al. Use of Telehealth by Surgical Specialties During the COVID-19 Pandemic. JAMA Surg 2021 Jul 01;156(7):620-626. [doi: 10.1001/jamasurg.2021.0979] [Medline: 33769434]
3. Koonin M, Hoots B, Tsang C. Trends in the use of telehealth during the emergence of the COVID-19 pandemic — United States, January–March 2020. MMWR Morb Mortal Wkly Rep 2020 Nov 13;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6945a9] [Medline: 33180760]
4. Verma S. Early Impact Of CMS Expansion Of Medicare Telehealth During COVID-19. Health Affairs Blog. 2020 Jul 15. URL: https://www.healthaffairs.org/do/10.1377/hblog20200715.454789/full/ [accessed 2020-08-01]
5. Billing for Telehealth During COVID-19. URL: https://tinyurl.com/2p9a5c84 [accessed 2021-08-01]
6. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field. J Am Med Inform Assoc 2020 Apr 23:1132-1135 [FREE Full text] [doi: 10.1093/jamia/ocaa072] [Medline: 32324855]
7. Gilson SF, Umscheid CA, Laiteerapong N, Ossey G, Nunes KJ, Shah SD. Growth of Ambulatory Virtual Visits and Differential Use by Patient Sociodemographics at One Urban Academic Medical Center During the COVID-19 Pandemic: Retrospective Analysis. JMIR Med Inform 2020 Dec 04;8(12):e24544-e24544 [FREE Full text] [doi: 10.2196/24544] [Medline: 33191247]
8. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. BMJ 2020 Mar 12;368:m998-m998. [doi: 10.1136/bmj.m998] [Medline: 32165352]
9. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 03;7(8):e016242-e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
10. Nguyen M, Waller M, Pandya A, Portnoy J. A Review of Patient and Provider Satisfaction with Telemedicine. Curr Allergy Asthma Rep 2020 Sep 22;20(11):72-72 [FREE Full text] [doi: 10.1007/s11882-020-00969-7] [Medline: 32959158]
11. Telehealth Interventions to Improve Chronic Disease. URL: https://www.cdc.gov/dhdsp/pubs/telehealth.htm [accessed 2021-08-01]
12. Agency for Healthcare Research and Quality. The Evidence Base for Telehealth: Reassurance in the Face of Rapid Expansion During the COVID-19 Pandemic. 2020 May 14. URL: https://effectivehealthcare.ahrq.gov/products/telehealth-expansion/white-paper [accessed 2021-08-01]
13. Doshi H, Hsia B, Shahani J, Mowrey W, Jariwala SP. Impact of Technology-Based Interventions on Patient-Reported Outcomes in Asthma: A Systematic Review. J Allergy Clin Immunol Pract 2021 Mar 04:2336-2341. [doi: 10.1016/j.jaip.2021.01.027] [Medline: 33548519]

14. Hanach N, de Vries N, Radwan H, Bissani N. The effectiveness of telemedicine interventions, delivered exclusively during the postnatal period, on postpartum depression in mothers without history or existing mental disorders: A systematic review and meta-analysis. Midwifery 2021 Mar;94:102906-102906 [FREE Full text] [doi: 10.1016/j.midw.2020.102906] [Medline: 33360589]
15. Hubley S, Lynch SB, Schneck C, Thomas M, Shore J. Review of key telepsychiatry outcomes. World J Psychiatry 2016 Jun 22;6(2):269-282 [FREE Full text] [doi: 10.5498/wjp.v6.i2.269] [Medline: 27354970]
16. Kraef C, van der Meirschen M, Free C. Digital telemedicine interventions for patients with multimorbidity: a systematic review and meta-analysis. BMJ Open 2020 Oct 13;10(10):e036904-e036904 [FREE Full text] [doi: 10.1136/bmjopen-2020-036904] [Medline: 33051232]
17. Noel H, Vogel D, Erdos J, Cornwall D, Levin F. Home telehealth reduces healthcare costs. Telemed J E Health 2004;10(2):170-183 [FREE Full text] [doi: 10.1089/tmj.2004.10.170] [Medline: 15319047]
18. Zhang W, Cheng B, Zhu W, Huang X, Shen C. Effect of Telemedicine on Quality of Care in Patients with Coexisting Hypertension and Diabetes: A Systematic Review and Meta-Analysis. Telemed J E Health 2021 Jun;27(6):603-614. [doi: 10.1089/tmj.2020.0122] [Medline: 32976084]
19. Sülz S, van Elten HJ, Askari M, Weggelaar-Jansen AM, Huijsman R. eHealth Applications to Support Independent Living of Older Persons: Scoping Review of Costs and Benefits Identified in Economic Evaluations. J Med Internet Res 2021 Mar 09;23(3):e24363-e24363 [FREE Full text] [doi: 10.2196/24363] [Medline: 33687335]
20. Ramsetty A, Adams C. Impact of the digital divide in the age of COVID-19. J Am Med Inform Assoc 2020 Apr 28:1147-1148 [FREE Full text] [doi: 10.1093/jamia/ocaa078] [Medline: 32343813]
21. Mitchell U, Chebli P, Ruggiero L, Muramatsu N. The Digital Divide in Health-Related Technology Use: The Significance of Race/Ethnicity. Gerontologist 2019 Jan 09;59(1):6-14. [doi: 10.1093/geront/gny138] [Medline: 30452660]
22. Cornejo Müller A, Wachtler B, Lampert T. [Digital divide-social inequalities in the utilisation of digital healthcare]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2020 Feb;63(2):185-191 [FREE Full text] [doi: 10.1007/s00103-019-03081-y] [Medline: 31915863]
23. Nouri S, Khoong E, Lyles C. Addressing Equity in Telemedicine for Chronic Disease Management During the Covid-19 Pandemic. NEJM Catalyst 2020 May:1-13 [FREE Full text] [doi: 10.1056/CAT.20.0123]
24. Annaswamy TM, Verduzco-Gutierrez M, Frieden L. Telemedicine barriers and challenges for persons with disabilities: COVID-19 and beyond. Disabil Health J 2020 Oct;13(4):100973-100973 [FREE Full text] [doi: 10.1016/j.dhjo.2020.100973] [Medline: 32703737]
25. Rodriguez JA, Betancourt JR, Sequist TD, Ganguli I. Differences in the use of telephone and video telemedicine visits during the COVID-19 pandemic. Am J Manag Care 2021 Jan;27(1):21-26 [FREE Full text] [doi: 10.37765/ajmc.2021.88573] [Medline: 33471458]
26. Schifeling CH, Shanbhag P, Johnson A, Atwater RC, Koljack C, Parnes BL, et al. Disparities in Video and Telephone Visits Among Older Adults During the COVID-19 Pandemic: Cross-Sectional Analysis. JMIR Aging 2020 Nov 10;3(2):e23176-e23176 [FREE Full text] [doi: 10.2196/23176] [Medline: 33048821]
27. Jaklevic MC. Telephone visits surge during the pandemic, but will they last? JAMA 2020 Oct 27;324(16):1593-1595. [doi: 10.1001/jama.2020.17201] [Medline: 33027503]
28. Uscher-Pines L, Sousa J, Jones M, Whaley C, Perrone C, McCullough C, et al. Telehealth Use Among Safety-Net Organizations in California During the COVID-19 Pandemic. JAMA 2021 Mar 16;325(11):1106-1107 [FREE Full text] [doi: 10.1001/jama.2021.0282] [Medline: 33528494]
29. Shah S, Alkureishi L, Lee W. Low-tech but high-value: Don't stop reimbursing phone-only telehealth visits. STAT News. URL: https://www.statnews.com/2021/04/21/low-tech-but-high-value-dont-stop-reimbursing-phone-only-telehealth-visits/ [accessed 2021-08-01]
30. American Medical Association. Why Audio-Only Telehealth Visits Must Continue. URL: https://tinyurl.com/874cc4et [accessed 2021-08-05]
31. DoBias M. Improving Equitable Access to Care for Seniors: Congress, CMS Need to Quickly Act on Audio-Only Risk Adjustment Bipartisan legislation in the House and Senate would create equity in the Medicare Advantage telehealth benefit. Alliance of Community Health Plans. URL: https://achp.org/equitable-access-through-risk-adjustment/?utm_source=Twitter&utm_medium=social&utm_campaign=brand_and_staff [accessed 2021-08-01]
32. Donaghy E, Atherton H, Hammersley V, McNeilly H, Bikker A, Robbins L, et al. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. Br J Gen Pract 2019 Sep;69(686):e586-e594 [FREE Full text] [doi: 10.3399/bjgp19X704141] [Medline: 31160368]
33. Gordon HS, Solanki P, Bokhour BG, Gopal RK. "I'm Not Feeling Like I'm Part of the Conversation" Patients' Perspectives on Communicating in Clinical Video Telehealth Visits. J Gen Intern Med 2020 Jun;35(6):1751-1758. [doi: 10.1007/s11606-020-05673-w] [Medline: 32016705]
34. Bleyel C, Hoffmann M, Wensing M, Hartmann M, Friederich H, Haun MW. Patients' Perspective on Mental Health Specialist Video Consultations in Primary Care: Qualitative Preimplementation Study of Anticipated Benefits and Barriers. J Med Internet Res 2020 Apr 20;22(4):e17330 [FREE Full text] [doi: 10.2196/17330] [Medline: 32310139]

35. Powell RE, Henstenburg JM, Cooper G, Hollander JE, Rising KL. Patient Perceptions of Telehealth Primary Care Video Visits. Ann Fam Med 2017 Dec;15(3):225-229 [FREE Full text] [doi: 10.1370/afm.2095] [Medline: 28483887]
36. Ray KN, Ashcraft LE, Mehrotra A, Miller E, Kahn JM. Family Perspectives on Telemedicine for Pediatric Subspecialty Care. Telemed J E Health 2017 Oct;23(10):852-862 [FREE Full text] [doi: 10.1089/tmj.2016.0236] [Medline: 28430021]
37. Sauers-Ford HS, Hamline MY, Gosdin MM, Kair LR, Weinberg GM, Marcin JP, et al. Acceptability, Usability, and Effectiveness: A Qualitative Study Evaluating a Pediatric Telemedicine Program. Acad Emerg Med 2019 Sep;26(9):1022-1033 [FREE Full text] [doi: 10.1111/acem.13763] [Medline: 30974004]
38. ATLAS.ti. URL: https://atlasti.com/ [accessed 2020-09-01]
39. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
40. World Health Organization. 10 Facts on Dementia. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2021-09-02]
41. Laver K, Liu E, Clemson L, Davies O, Gray L, Gitlin LN, et al. Does Telehealth Delivery of a Dyadic Dementia Care Program Provide a Noninferior Alternative to Face-To-Face Delivery of the Same Program? A Randomized, Controlled Trial. Am J Geriatr Psychiatry 2020 Jun;28(6):673-682 [FREE Full text] [doi: 10.1016/j.jagp.2020.02.009] [Medline: 32234275]
42. Sakaguchi-Tang D, Bosold A, Choi YK, Turner AM. Patient Portal Use and Experience Among Older Adults: Systematic Review. JMIR Med Inform 2017 Oct 16;5(4):e38 [FREE Full text] [doi: 10.2196/medinform.8092] [Medline: 29038093]
43. Nissen R, Hersch G, Tietze M. Persons With Dementia and Their Caregivers' Perceptions About Occupational Therapy and Telehealth: A Qualitative Descriptive Study. Home Healthc Now. Nov/Dec 2018;36(6):A. [doi: 10.1097/nhh.0000000000000697]
44. Alison B. World Failing to Address Dementia Challenge. Sept 2, 2021.: World Health Organization URL: https://www.who.int/news/item/02-09-2021-world-failing-to-address-dementia-challenge [accessed 2021-09-02]
45. Zajacova A, Lawrence EM. The Relationship Between Education and Health: Reducing Disparities Through a Contextual Approach. Annu. Rev. Public Health 2018 Apr;39(1):273-289. [doi: 10.1146/annurev-publhealth-031816-044628]
46. The Homework Gap. National School Boards Association. URL: https://www.nsba.org/Advocacy/Federal-Legislative-Priorities/Homework-Gap [accessed 2021-08-01]
47. McLaughlin C. The Homework Gap: The Cruelest Part of the Digital Divide. National Education Association. 2016 Apr 20. URL: https://www.nea.org/advocating-for-change/new-from-nea/homework-gap-cruelest-part-digital-divide [accessed 2021-08-01]
48. McElrath K. Schooling During the COVID-19 Pandemic. United States Census Bureau. URL: https://www.census.gov/library/stories/2020/08/schooling-during-the-covid-19-pandemic.html [accessed 2021-08-01]
49. Common SM. The U.S. K-12 Digital Divide Has Narrowed, but Must Close to Eliminate Risks to Students and the Economy. Common Sense, and the Southern Education Foundation. 2021 Jan 26. URL: https://www.commonsensemedia.org/about-us/news/press-releases/the-us-k-12-digital-divide-has-narrowed-but-must-close-to-eliminate [accessed 2021-08-01]
50. Hall L. Digital Redlining. Institute for Emerging Issues. 2020 Aug 14. URL: https://iei.ncsu.edu/2020/digital-redlining/ [accessed 2021-08-01]
51. Blanchard J, Blanchard A. Slow internet? How digital redlining hurts people of all ages. Chicago Tribune. 2020 Jul 24. URL: https://www.chicagotribune.com/opinion/commentary/ct-opinion-coronavirus-isolation-internet-connection-20200724-2ankrcnvorgpbglu5p53e7etgy-story.html [accessed 2020-07-24]
52. Ali T, Chandra S, Cherukumilli S. Looking Back, Looking Forward: What it will take to permanently close the K-12 digital divide. Benton Institute. 2021 Jan 27. URL: https://www.benton.org/headlines/looking-back-looking-forward-what-it-will-take-permanently-close-k-12-digital-divide [accessed 2021-08-01]
53. Gibbons MC. A historical overview of health disparities and the potential of eHealth solutions. J Med Internet Res 2005;7(5):e50 [FREE Full text] [doi: 10.2196/jmir.7.5.e50] [Medline: 16403714]
54. Marcin JP, Shaikh U, Steinhorn RH. Addressing health disparities in rural communities using telehealth. Pediatr Res 2016 Jan;79(1-2):169-176. [doi: 10.1038/pr.2015.192] [Medline: 26466080]
55. Talal AH, Sofikitou EM, Jaanimägi U, Zeremski M, Tobin JN, Markatou M. A framework for patient-centered telemedicine: Application and lessons learned from vulnerable populations. J Biomed Inform 2020 Dec;112:103622. [doi: 10.1016/j.jbi.2020.103622] [Medline: 33186707]
56. Bakhtiar M, Elbuluk N, Lipoff JB. The digital divide: How COVID-19's telemedicine expansion could exacerbate disparities. Journal of the American Academy of Dermatology 2020 Nov;83(5):e345-e346. [doi: 10.1016/j.jaad.2020.07.043]
57. Watts G. COVID-19 and the digital divide in the UK. The Lancet Digital Health 2020 Aug;2(8):e395-e396. [doi: 10.1016/s2589-7500(20)30169-2]
58. Scanzera AC, Kim SJ, Paul Chan RV. Teleophthalmology and the digital divide: inequities highlighted by the COVID-19 pandemic. Eye 2020 Nov 30;35(6):1529-1531. [doi: 10.1038/s41433-020-01323-x]

59. Jennifer B. AMIA: Consider Broadband Access a Social Determinant of Health. Health IT Analytics. 2017 May 25. URL: https://healthitanalytics.com/news/amia-consider-broadband-access-a-social-determinate-of-health [accessed 2021-08-01]
60. Healthy People 2030, Social Determinants of Health. U.S. Department of Health and Human Services. URL: https://health.gov/healthypeople/objectives-and-data/social-determinants-health [accessed 2021-08-01]
61. About Social Determinants of Health (SDOH). Centers for Disease Control and Prevention.: Centers for Disease Control and Prevention URL: https://www.cdc.gov/socialdeterminants/about.html [accessed 2021-07-31]
62. Press V, Huisingh-Scheetz M, Arora V. Op-ed: COVID-19 vaccine registration by app? Confusing technology reveals challenges for seniors using telemedicine. Chicago Tribune. URL: https://www.chicagotribune.com/opinion/commentary/ct-opinion-vaccine-seniors-digital-literacy-telemedicine-20210120-x6mbmd2swjdanouifuqkfmll2q-story.html [accessed 2021-02-01]
63. Arend J, Tsang-Quinn J, Levine C, Thomas D. The Patient-Centered Medical Home: History, Components, and Review of the Evidence. Mt Sinai J Med 2012 Jul 11;79(4):433-450. [doi: 10.1002/msj.21326]
64. Kraschnewski JL, Gabbay RA. Role of Health Information Technologies in the Patient-Centered Medical Home. J Diabetes Sci Technol 2013 Sep 01;7(5):1376-1385. [doi: 10.1177/193229681300700530]
65. Kang C. Parking Lots Have Become a Digital Lifeline. The New York Times. 2020 May 20. URL: https://www.nytimes.com/2020/05/05/technology/parking-lots-wifi-coronavirus.html [accessed 2021-08-01]
66. Egan P. Map shows free Wi-Fi hotspots in Michigan: How to find one near you. Detroit Free Press. 2020 Jun 10. URL: https://www.freep.com/story/news/local/michigan/2020/06/10/free-wifi-hotspot-internet-map-michigan/5332565002/ [accessed 2021-08-01]
67. WiFi Hotspots. Connected Nation.: Connected Nation URL: https://cngis.maps.arcgis.com/apps/webappviewer/index.html?id=0d69accbb5ff422a82eccc2c9101b69d [accessed 2021-08-01]
68. Blanchard J, Blanchard A. Slow internet? How digital redlining hurts people of all ages. Chicago Tribune. 2020 Jul 24. URL: https://www.chicagotribune.com/opinion/commentary/ct-opinion-coronavirus-isolation-internet-connection-20200724-2ankrcnvorgpbglu5p53e7etgy-story.html [accessed 2021-08-01]
69. Levin B, Downes L. Cities, not rural areas, are the real Internet deserts. New York Times. 2019 Sep 13. URL: https://www.washingtonpost.com/technology/2019/09/13/cities-not-rural-areas-are-real-internet-deserts/ [accessed 2021-08-01]
70. Denson J. A Parched Country: Study shows digital deserts exist. Connected Nation. 2019 Sep 18. URL: https://connectednation.org/blog/2019/09/18/a-parched-country-study-shows-digital-deserts-exist/ [accessed 2021-08-01]
71. Federal CC. 2020 Broadband Deployment Report. Federal Communications Commission.: Federal Communications Commission; 2020 Jun 08. URL: https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2020-broadband-deployment-report) [accessed 2020-08-01]
72. Hackett M. Designing better tech for seniors means simplifying tech for everyone. MobiHealth News. 2020 Nov. URL: https://www.mobihealthnews.com/news/designing-better-tech-seniors-means-simplifying-tech-everyone [accessed 2021-08-01]
73. Improving Literacy in the United States: Recommendations for Increasing Reading Success. The Center for American Progress.: The Center for American Progress; 2020 May 28. URL: https://www.americanprogress.org/issues/education-k [accessed 2020-08-23]
74. Howard J. The complicated role of the modern public library. Humanities: The Magazine of the National Endowment for the Humanities. URL: https://www.neh.gov/article/complicated-role-modern-public-library [accessed 2021-08-01]
75. Klinenberg E. Palaces for the People: How Social Infrastructure Can Help Fight Inequality, Polarization, and the Decline of Civic Life. New York, NY: Crown Publishing Group; 2018:1524761168.
76. Fallows D. The Post-pandemic Future of Libraries. The Atlantic. 2020 May 12. URL: https://www.theatlantic.com/notes/2020/05/post-pandemic-future-libraries/611458/ [accessed 2021-08-01]
77. Fallows D. Public Libraries? Novel Response to a Novel Virus. The Atlantic. 2020 Mar 31. URL: https://www.theatlantic.com/notes/2020/03/public-libraries-novel-response-to-a-novel-virus/609058/ [accessed 2021-08-01]
78. Guernsey L, Prescott S, Park C. Public Libraries and the Pandemic: Digital Shifts and Disparities to Overcome. Education Policy Open Technology Institute. 2021 Mar. URL: https://www.newamerica.org/education-policy/reports/public-libraries-and-the-pandemic/ [accessed 2021-08-01]
79. Guernsey L. Library Survey Data Shows Internet Gaps and Need for Federal Action. New America. 2021 Mar. URL: https://www.newamerica.org/education-policy/edcentral/library-survey-data-shows-internet-gaps-and-need-for-federal-action/ [accessed 2021-08-01]
80. NowPow. URL: https://nowpow.com/ [accessed 2021-08-01]
81. Humana Partners With NowPow To Address Social Needs. Business Wire. 2021 May 18. URL: https://www.businesswire.com/news/home/20210518005306/en/Humana-Partners-With-NowPow-To-Address-Social-Needs [accessed 2021-08-01]
82. Press VG, Huisingh-Scheetz M, Arora VM. Inequities in Technology Contribute to Disparities in COVID-19 Vaccine Distribution. JAMA Health Forum 2021 Mar 19;2(3):e210264. [doi: 10.1001/jamahealthforum.2021.0264]

83. Hackett M. Eliminating audio-only telehealth coverage could put safety net clinics at risk: CMS has signaled that it may stop reimbursing for audio-only visits after the public health emergency. Healthcare Finance. 2021 Feb. URL: https://www.healthcarefinancenews.com/news/eliminating-audio-only-telehealth-coverage-could-put-safety-net-clinics-risk [accessed 2021-08-01]
84. Baras Shreibati J. When Low Tech Wins. N Engl J Med 2021 Aug 12;385(7):581-583 [FREE Full text] [doi: 10.1056/nejmp2104234]
85. Uscher-Pines L, Jones M, Sousa J. The doctor will call me maybe: the uncertain future of audio-only visits and why we need them to address disparities. Health Affairs Blog. 2021 Mar 03. URL: https://www.healthaffairs.org/do/10.1377/hblog20210225.26462/full [accessed 2021-07-31]
86. Mehrotra A, Bhatia RS, Snoswell CL. Paying for Telemedicine After the Pandemic. JAMA 2021 Feb 02;325(5):431-432. [doi: 10.1001/jama.2020.25706] [Medline: 33528545]
87. Czeisler, Marynak K, Clarke KE, Salah Z, Shakya I, Thierry JM, et al. Delay or Avoidance of Medical Care Because of COVID-19–Related Concerns — United States, June 2020. MMWR Morb. Mortal. Wkly. Rep 2020 Sep 11;69(36):1250-1257. [doi: 10.15585/mmwr.mm6936a4]
88. Findling MG, Blendon RJ, Benson JM. Delayed Care with Harmful Health Consequences—Reported Experiences from National Surveys During Coronavirus Disease 2019. JAMA Health Forum 2020 Dec 14;1(12):e201463. [doi: 10.1001/jamahealthforum.2020.1463]
89. Sperber NR, King HA, Steinhauser K, Ammarell N, Danus S, Powers BJ. Scheduled telephone visits in the veterans health administration patient-centered medical home. BMC Health Serv Res 2014 Apr 01;14:145 [FREE Full text] [doi: 10.1186/1472-6963-14-145] [Medline: 24690086]
90. Bernstein P, Ko KJ, Israni J, Cronin AO, Kurliand MM, Shi JM, et al. Urgent and non-emergent telehealth care for seniors: Findings from a multi-site impact study. J Telemed Telecare 2021 Apr 17:1357633X211004321 [FREE Full text] [doi: 10.1177/1357633X211004321] [Medline: 33866894]
91. Irvine A, Drew P, Bower P, Brooks H, Gellatly J, Armitage CJ, et al. Are there interactional differences between telephone and face-to-face psychological therapy? A systematic review of comparative studies. J Affect Disord 2020 Mar 15;265:120-131 [FREE Full text] [doi: 10.1016/j.jad.2020.01.057] [Medline: 32090733]
92. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. BMC Med Res Methodol 2014;14(1):42. [doi: 10.1186/1471-2288-14-42]

Abbreviations

CMS: Centers for Medicare and Medicaid Services
SDOH: social determinants of health
UCMC: University of Chicago Medical Center

Edited by G Eysenbach; submitted 07.09.21; peer-reviewed by E van der Velde, K Laver, E Donaghy; comments to author 23.09.21; revised version received 14.10.21; accepted 24.10.21; published 15.12.21.

Please cite as:
Alkureishi MA, Choo ZY, Rahman A, Ho K, Benning-Shorb J, Lenti G, Velázquez Sánchez I, Zhu M, Shah SD, Lee WW
Digitally Disconnected: Qualitative Study of Patient Perspectives on the Digital Divide and Potential Solutions
JMIR Hum Factors 2021;8(4):e33364
URL: https://humanfactors.jmir.org/2021/4/e33364
doi:10.2196/33364
PMID:34705664

Original Paper

Communicating Guideline Recommendations Using Graphic Narrative Versus Text-Based Broadcast Screensavers: Design and Implementation Study

Lauren Sinnenberg[1], MD; Craig A Umscheid[2,3], MSc, MD; Frances S Shofer[4,5], PhD; Damien Leri[6], MEd; Zachary F Meisel[4,5], MPH, MSHP, MD

[1]Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States
[2]Biological Sciences Division, University of Chicago, Chicago, IL, United States
[3]Center for Healthcare Delivery Science and Innovation, University of Chicago Medicine, Chicago, IL, United States
[4]Department of Emergency Medicine, University of Pennsylvania, Philadelphia, PA, United States
[5]Center for Emergency Care Policy and Research, University of Pennsylvania, Philadelphia, PA, United States
[6]Penn Medicine Center for Health Care Innovation, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Zachary F Meisel, MPH, MSHP, MD
Center for Emergency Care Policy and Research
University of Pennsylvania
Ravdin Ground
3400 Spruce Street
Philadelphia, PA, 19104
United States
Phone: 1 215 746 5618
Email: zfm@pennmedicine.upenn.edu

Abstract

Background: The use of graphic narratives, defined as stories that use images for narration, is growing in health communication.

Objective: The aim of this study was to describe the design and implementation of a graphic narrative screensaver (GNS) to communicate a guideline recommendation (ie, avoiding low-value acid suppressive therapy [AST] use in hospital inpatients) and examine the comparative effectiveness of the GNS versus a text-based screensaver (TBS) on clinical practice (ie, low-value AST prescriptions) and clinician recall.

Methods: During a 2-year period, the GNS and the TBS were displayed on inpatient clinical workstations. The numbers of new AST prescriptions were examined in the four quarters before, the three quarters during, and the one quarter after screensavers were implemented. Additionally, an electronic survey was sent to resident physicians 1 year after the intervention to assess screensaver recall.

Results: Designing an aesthetically engaging graphic that could be rapidly understood was critical in the development of the GNS. The odds of receiving an AST prescription on medicine and medicine subspecialty services after the screensavers were implemented were lower for all four quarters (ie, GNS and TBS broadcast together, only TBS broadcast, only GNS broadcast, and no AST screensavers broadcast) compared to the quarter prior to implementation (odds ratio [OR] 0.85, 95% CI 0.78-0.92; OR 0.89, 95% CI 0.82-0.97; OR 0.87, 95% CI 0.80-0.95; and OR 0.81, 95% CI 0.75-0.89, respectively; P<.001 for all comparisons). There were no statistically significant decreases for other high-volume services, such as the surgical services. These declines appear to have begun prior to screensaver implementation. When surveyed about the screensaver content 1 year later, resident physicians recalled both the GNS and TBS (43/70, 61%, vs 54/70, 77%; P=.07) and those who recalled the screensaver were more likely to recall the main message of the GNS compared to the TBS (30/43, 70%, vs 1/54, 2%; P<.001).

Conclusions: It is feasible to use a graphic narrative embedded in a broadcast screensaver to communicate a guideline recommendation, but further study is needed to determine the impact of graphic narratives on clinical practice.

(JMIR Hum Factors 2021;8(4):e27171) doi:10.2196/27171

KEYWORDS

medical informatics; screensaver; guideline dissemination; graphic narratives; health communication; workstation; clinical workstation; guidelines; medical education; education

Introduction

The use of graphic narratives is growing in health communication [1]. They are characterized as cohesive stories with an identifiable beginning, middle, and end that include characters, raise questions, provide resolution, and use images for narration. Graphic narratives have been successfully used by the American Cancer Society and the US Centers for Disease Control and Prevention in patient-facing communication [2,3]. The theoretical underpinnings for behavior change resulting from narratives include social cognitive theory and the theory of reasoned action [4].

Prior work in health communication has also evaluated the use of broadcast screensavers as educational tools for disseminating information to hospital staff, with mixed results [5-7]. No study to date has specifically examined the comparative effectiveness of different approaches to communicate messages using broadcast screensavers targeted to health care providers.

In this study, we describe the feasibility of designing and implementing a graphic narrative to communicate a guideline recommendation—namely, avoiding acid suppressive therapy (AST) in hospital inpatients at low risk of gastric stress ulcers—to health care providers through the use of broadcast screensavers, and we examine the comparative effectiveness of a graphic narrative screensaver (GBS) versus a text-based screensaver (TBS) on clinical practice and clinician recall [8-11].

Methods

Overview

This was a descriptive feasibility study as well as a quasi-experimental evaluative study that examined change in clinical practice over a 2-year period and included an experimental survey component to examine clinician recall. The study site was a single academic health care system consisting of three hospitals in an urban environment. The study was approved by the Institutional Review Board of the University of Pennsylvania.

Graphic Narrative Design

The GNS and the TBS were designed to communicate the risk of unindicated AST prescription. We developed narratives through meetings in which feedback around the low-value prescription of AST was solicited in a semistructured manner from health care system faculty, nurses, fellows, and residents. The focus of these meetings was to elicit knowledge gaps, attitudes, and beliefs related to AST use. We contracted a graphic designer to create a slide that could be broadcast on the screensaver of inpatient clinical workstations and used established techniques and theoretical frameworks in narrative communication [4]. The slide was developed and refined in an iterative fashion in which the designer presented ideas and prototypes in three rounds to the research team, which was composed of decision scientists and clinicians. We simultaneously developed text-based, probabilistic descriptions of the published guidelines from the Choosing Wisely campaign to compare with the graphic narratives.

Screensaver Intervention

Study screensavers were added to an existing, rotating deck of screensaver slides updated on the first of each month and displayed on all clinical workstation computers of all inpatient units in the three urban hospitals of our academic health care system. Slides were displayed from the deck in random order, lasting 18 seconds per slide. In most months, there are 10 or fewer slides in rotation, and rarely are there more than 20 slides in rotation. For the initial 3-month block of our study intervention period (October to December 2014), the GNS and TBS were both included in the slide deck for broadcasting. For the next 3 months (January to March 2015), only the TBS was included in the slide deck for broadcasting, followed by a 3-month block (April to June 2015) where only the GNS was included in the slide deck. The final 3 months of the study period (July to September 2015) included neither of the study screensavers.

AST Prescriptions

New discharge prescriptions of AST for all low-risk inpatients were measured prior to, during, and following implementation of the intervention screensavers using data from our health care system's electronic medical record (Allscripts). Patients were included if they were admitted and discharged during our study period of October 1, 2013, to September 30, 2015. Inpatients were defined as "low risk" using criteria from the American Society of Health-System Pharmacists guideline on AST use [12]. To ensure we included only low-risk inpatients in our analysis, the following patients were excluded: intensive care unit patients with an international normalized ratio of >1.9 or partial thromboplastin time of >54, patients on mechanical ventilation, patients with a history of or current peptic ulcer disease (ICD-9 [International Classification of Diseases, Ninth Revision] codes 531-533), and patients cared for on the clinical research unit, hospice, and gastroenterology medical or surgical service. Patients who were less than 18 years of age, left against medical advice, expired during hospitalization, or discharged to hospice were also excluded. AST was defined as any of the following: proton pump inhibitors (ie, lansoprazole, omeprazole, pantoprazole, dexlansoprazole, esomeprazole, and rabeprazole) or histamine H2-receptor antagonists (ie, cimetidine, famotidine, nizatidine, and ranitidine).

Resident Physician Survey

One year after the screensavers were broadcast, an electronic questionnaire was emailed to all second- and third-year internal medicine resident physicians. First-year residents were excluded as they had not been exposed to the intervention.

To evaluate retention of guideline information, the survey had an experimental design. Participants were shown the slides from the screensavers with all written content deliberately blurred.

Residents were then asked to recall if they had seen the slide and to describe the content of the slide from memory. Participants were also asked about prescribing patterns of AST, adverse effects of AST, and how their prescribing patterns had changed in the previous year. The survey was emailed four times to the study population. To compensate the residents for their participation, they were entered into a lottery for one of six US $25 gift cards or an Apple Watch.

Statistical Analysis

Summary statistics, such as frequencies and percentages or means and SDs, were used to describe the patient population in terms of sex, race, age, hospital, and clinical service. Clinical discharge service was divided into six groups: (1) medicine (eg, general internal medicine), (2) surgery (eg, orthopedics, neurosurgery, general surgery, and urology), (3) medical subspecialty services (eg, oncology, cardiology, pulmonary, and infectious disease), (4) family medicine, (5) neurology, and (6) obstetrics and gynecology. The 2-year study period was divided into eight, 3-month blocks: prequarter 1, prequarter 2, prequarter 3, prequarter 4, GNS and TBS, TBS alone, GNS alone, and postquarter. To assess trends in AST prescriptions, logistic regressions that were modeled on receiving a new AST prescription on discharge, adjusted for sex, race, age, and hospital, were developed. The models for each 3-month period were compared to prequarter 4 (ie, the 3-month period prior to the intervention).

Standard summary statistics were used to describe participants in the survey. To compare differences in screensaver recall, chi-square and McNemar tests were used. All analyses were performed using SAS statistical software (version 9.4; SAS Institute).

Results

Graphic Narrative Design and Implementation

During our semistructured meetings, we found the following to be important features of graphic narrative design: crafting the guideline-based message into narrative form; creating a narrative that is attention grabbing, such that it attracts busy hospital staff; and ensuring that the graphic narrative is quickly comprehensible. We contracted a designer who was able to create a graphic narrative that met these key requirements (Figure 1).

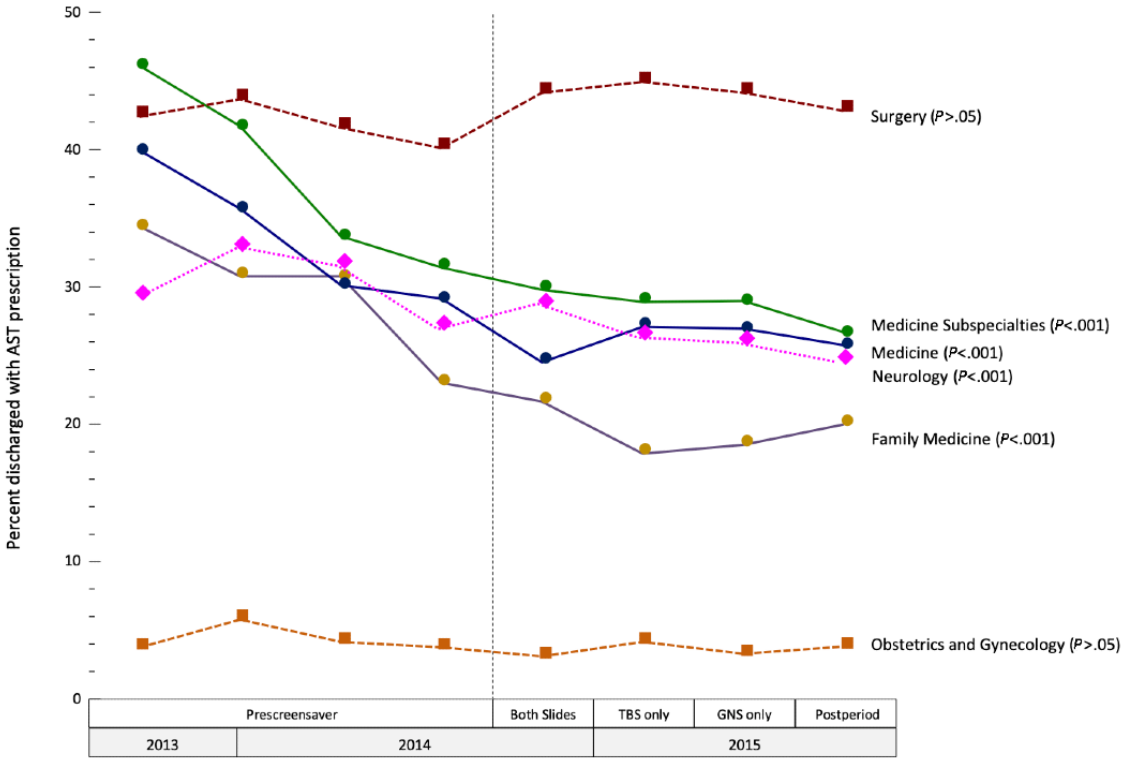
Figure 1. Screensaver interventions (A) and experimental survey design (B) containing screensavers blinded by blurring all content-specific text.

Effect of GNS and TBS on Acid Suppressive Therapy Prescriptions

During the 2-year period, 157,110 patients were admitted to one of the three study hospitals, of which 97,767 met the inclusion criteria (62.2%). The patient sample was 60.9% (n=59,495) male, 53.3% (n=51,017) White, and 41.1% (n=39,419) African American or Black, and had a mean age of 52.0 (SD 19.2) years. Most patients were discharged from surgical services (n=31,429, 32.1%), followed by obstetrics and gynecology (n=21,117, 21.6%), internal medicine (n=20,934, 21.4%), internal medicine subspecialty (n=20,592, 21.1%), neurology (n=2526, 2.6%), and family medicine (n=1169, 1.2%) services. A total of 56.0% (n=54,799) of the patients were discharged directly home.

After adjusting for sex, race, age, and hospital, for both medicine and medicine subspecialty services combined, the odds of receiving an AST prescription after the screensaver interventions were implemented was lower for all four quarters (ie, GNS and TBS, TBS alone, GNS alone, and postquarter) compared to prequarter 4 (odds ratio [OR] 0.85, 95% CI 0.78-0.92; OR 0.89, 95% CI 0.82-0.97; OR 0.87, 95% CI 0.80-0.95; and OR 0.81, 95% CI 0.75-0.89, respectively; P<.001 for all comparisons). There were no statistically significant decreases for the other services. These declines appear to have begun prior to screensaver implementation (Figure 2).

Figure 2. Acid suppressive therapy (AST) prescription patterns during the study period. GNS: graphic narrative screensaver; TBS: text-based screensaver.



Resident Physician Survey

Of the 97 residents invited to participate, 70 (72%) completed the survey. The median age of participants was 29 (IQR 2) years, and 51% (n=36) were male. Most residents indicated that they could recall seeing both the GNS and TBS (n=43, 61%, vs n=54, 77%; P=.07). When those who recalled seeing the screensavers were asked where they had seen the image, 93% (40/43) recalled that the GNS was a screensaver, compared to 24% (13/54) for the TBS. Furthermore, 70% (30/43) could recall the main topic of the GNS, compared to 2% (1/54) of the TBS (P<.001). Many residents indicated that they prescribed fewer ASTs than they did 1 year prior (38/70, 54%), and 8% (3/38) of these participants directly attributed their change to the screensavers.

Discussion

We sought to design a GNS that communicated guideline recommendations. In a design process that included semistructured meetings with key stakeholders as well as the efforts of a professional graphic designer, we found that it was feasible to create and disseminate a graphic narrative to summarize and communicate guideline recommendations.

In our study period, approximately one-quarter of patients were discharged with inappropriate AST prescriptions, but these AST prescriptions decreased over time on the nonsurgical services. This decrease, however, appeared to have begun prior to the screensaver intervention. It is possible that there were ongoing efforts on the nonsurgical services to reduce unnecessary AST prescriptions, and it is unknown whether the screensaver initiative may have potentiated this effect. The intervention seems to have had a lower effect on the surgical services, potentially related to less time spent by surgical service residents on the computer workstation.

Our study raises the possibility that GNSs may be useful tools for disseminating guideline recommendations. It is possible that the residents recognized the Choosing Wisely logo in the TBS leading to improved recognition compared to the GNS, although this difference was not statistically significant. This did not result in improved content-specific recall, however, which was significantly greater for the GNS compared to the TBS. This is consistent with prior work that demonstrated improved information delivery to clinicians when content was presented in narrative form as opposed to a summary statement form [13]. Further work should be done to understand whether graphic narratives have an impact on clinical practice and, if so, what features of graphic narratives improve information delivery.

Our study has limitations. We incorporated our screensavers into an existing broadcast screensaver program within an academic health care system. Guideline dissemination via broadcast screensavers may prove more challenging in

nonacademic settings without established broadcast screensaver programs. In addition, our quasi-experimental design prevents us from isolating the effects of our intervention from other interventions that may have concurrently affected AST prescription rates. The survey portion of our work had a relatively small sample size, which may have limited our ability to detect differences in recognition and recall. Future work should employ a randomized design to best isolate the effect of GNSs from other interventions.

In conclusion, it is feasible to use a graphic narrative embedded in a broadcast screensaver to communicate a guideline recommendation, but further study is needed to determine the impact of graphic narratives on clinical practice.

Acknowledgments

This work was supported by a Penn Implementation Science Working Group Pilot Grant.

Conflicts of Interest

None declared.

References

1. Green MJ, Myers KR. Graphic medicine: Use of comics in medical education and patient care. BMJ 2010 Mar 03;340:c863. [doi: 10.1136/bmj.c863] [Medline: 20200064]
2. Krakow M. Graphic narratives and cancer prevention: A case study of an American Cancer Society comic book. Health Commun 2017 May;32(5):525-528. [doi: 10.1080/10410236.2016.1211075] [Medline: 27542072]
3. Zombie preparedness graphic novel. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/cpr/zombie/novel.htm [accessed 2020-11-18]
4. Hinyard LJ, Kreuter MW. Using narrative communication as a tool for health behavior change: A conceptual, theoretical, and empirical overview. Health Educ Behav 2007 Oct;34(5):777-792. [doi: 10.1177/1090198106291963] [Medline: 17200094]
5. Cadena J, Prigmore T, Bowling J, Ayala BA, Kirkman L, Parekh A, et al. Improving influenza vaccination of healthcare workers by means of quality improvement tools. Infect Control Hosp Epidemiol 2011 Jun;32(6):616-618. [doi: 10.1086/660198] [Medline: 21558776]
6. Coil CJ, Kaji AH, Crevensten H, Aaron KE, Lewis RJ, Coates WC. Screen savers as an adjunct to medical education on patient safety. Jt Comm J Qual Patient Saf 2011 Nov;37(11):524-528, AP1-AP10. [doi: 10.1016/s1553-7250(11)37067-5]
7. Charters M, Cheng AL, Esaki RK, Kuo AS, Neal J, Thawani JP, et al. The impact of a screensaver in promoting hand hygiene. Am J Infect Control 2009 Dec;37(10):867-868. [doi: 10.1016/j.ajic.2009.05.003] [Medline: 20004813]
8. Society of Hospital Medicine – Adult hospital medicine: Five things physicians and patients should question. Choosing Wisely. 2013 Feb 21. URL: http://www.choosingwisely.org/societies/society-of-hospital-medicine-adult/ [accessed 2017-05-02]
9. Gupta R, Garg P, Kottoor R, Munoz JC, Jamal MM, Lambiase LR, et al. Overuse of acid suppression therapy in hospitalized patients. South Med J 2010;103(3):207-211. [doi: 10.1097/smj.0b013e3181ce0e7a]
10. Wohlt PD, Hansen LA, Fish JT. Inappropriate continuation of stress ulcer prophylactic therapy after discharge. Ann Pharmacother 2016 Nov 29;41(10):1611-1616. [doi: 10.1345/aph.1k227]
11. Albugeaey M, Alfaraj N, Garb J, Seiler A, Lagu T. Do hospitalists overuse proton pump inhibitors? Data from a contemporary cohort. J Hosp Med 2014 Nov;9(11):731-733 [FREE Full text] [doi: 10.1002/jhm.2249] [Medline: 25130756]
12. ASHP Commission on Therapeutics. ASHP therapeutic guidelines on stress ulcer prophylaxis. Am J Health Syst Pharm 1999 Feb 15;56(4):347-379. [doi: 10.1093/ajhp/56.4.347] [Medline: 10690219]
13. Meisel ZF, Metlay JP, Sinnenberg L, Kilaru AS, Grossestreuer A, Barg FK, et al. A randomized trial testing the effect of narrative vignettes versus guideline summaries on provider response to a professional organization clinical policy for safe opioid prescribing. Ann Emerg Med 2016 Dec;68(6):719-728 [FREE Full text] [doi: 10.1016/j.annemergmed.2016.03.007] [Medline: 27133392]

Abbreviations

AST: acid suppressive therapy
GNS: graphic narrative screensaver
ICD-9: International Classification of Diseases, Ninth Revision
TBS: text-based screensaver

Edited by A Kushniruk; submitted 14.01.21; peer-reviewed by JA Johnsen, M Murphy, S Sabarguna, L Brosseau; comments to author 28.03.21; revised version received 22.05.21; accepted 04.07.21; published 13.12.21.

Please cite as:
Sinnenberg L, Umscheid CA, Shofer FS, Leri D, Meisel ZF
Communicating Guideline Recommendations Using Graphic Narrative Versus Text-Based Broadcast Screensavers: Design and Implementation Study
JMIR Hum Factors 2021;8(4):e27171
URL: https://humanfactors.jmir.org/2021/4/e27171
doi:10.2196/27171
PMID:34264197

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org