JMIR Human Factors

Impact Factor (2022): 2.7
Volume 9 (2022), Issue 1 ISSN 2292-9495 Editor in Chief: Andre Kushniruk, BA, MSc, PhD, FACMI

Contents

Original Papers

Reviews

Letter to the Editor

Original Paper

A Novel Method for Digital Pain Assessment Using Abstract Animations: Human-Centered Design Approach

Nema Rao[1], MDes; Sophy Perdomo[2], MS, PhD; Charles Jonassaint[2], MHS, PhD

[1]Microsoft Inc, SharePoint Spaces, Seattle, WA, United States

[2]Center for Behavioral Health and Smart Technology, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Charles Jonassaint, MHS, PhD
Center for Behavioral Health and Smart Technology
Department of Medicine
University of Pittsburgh
230 McKee Pl Suite 600
Pittsburgh, PA, 15213
United States
Phone: 1 4125869850
Email: cjonassaint@pitt.edu

Abstract

Background: Patients with chronic pain face several challenges in using clinical tools to help them monitor, understand, and make meaningful decisions about their pain conditions. Our group previously presented data on Painimation, a novel electronic tool for communicating and assessing pain.

Objective: This paper describes the human-centered design and development approach (inspiration, ideation, and implementation) that led to the creation of Painimation.

Methods: We planned an iterative and cyclical development process that included stakeholder engagement and feedback from users. Stakeholders included patients with acute and chronic pain, health care providers, and design students. Target users were adults with acute or chronic pain who needed clinical assessment and tracking of the course of their pain over time. Phase I (inspiration) consisted of empathizing with users, understanding how patients experience pain, and identifying the barriers to accurately expressing and assessing pain. This phase involved understanding how patients communicate pain symptoms to providers, as well as defining limitations of current models of clinical pain assessment tools. In Phase II (ideate) we conceptualized and evaluated different approaches to expressing and assessing pain. The most promising concept was developed through an iterative process that involved end users and stakeholders. In Phase III (implementation), based on stakeholder feedback from initial designs and prototypes of abstract pain animations (painimations), we incorporated all concepts to test a minimally viable product, a fully functioning pain assessment app. We then gathered feedback through an agile development process and applied this feedback to finalizing a testable version of the app that could ultimately be used in a pain clinic.

Results: Engaging intended users and stakeholders in an iterative, human-centered design process identified 5 criteria that a pain assessment tool would need to meet to be effective in the medical setting. These criteria were used as guiding design principles to generate a series of pain assessment concept ideas. This human-centered approach generated 8 highly visual painimations that were found to be acceptable and useable for communicating pain with medical providers, by both patients with general pain and patients with sickle cell disease (SCD). While these initial steps continued refinement of the tool, further data are needed. Agile development will allow us to continue to incorporate precision medicine tools that are validated in the clinical research arena.

Conclusions: A multiphase, human-centered design approach successfully resulted in the development of an innovation that has potential to improve the quality of medical care, particularly for underserved populations. The use of Painimation may especially benefit the medical care of minority populations with chronic and difficult-to-treat pain, such as adults with SCD. The insights generated from this study can be applied to the development of patient-reported outcomes tools that are more patient-centered, engaging, and effective.

(JMIR Hum Factors 2022;9(1):e27689) doi:10.2196/27689

KEYWORDS

pain; pain measurement; chronic pain; animations; mobile apps; human-centered design

Introduction

Background

Pain is the number one reason people access the health care system. The Centers for Disease Control and Prevention (CDC) reported that in 2019, approximately 20.4% of US adults had chronic pain, and 7.4% had high-impact, chronic pain. Similar statistics have been reported in Canada (18.9%) and Australia (17.9%), whereas in the United Kingdom the numbers are much higher (35%-51.3%) [1-4]. The cost of medical treatment and lost productivity due to pain exceeds US $635 billion each year in the United States, more than the cost of treating cardiovascular disease, cancer, or diabetes [5]. Chronic pain also significantly affects an individual's quality of life, negatively impacting their ability to engage in day-to-day activities, and increasing risk for depression, anxiety, and opioid dependence [6,7].

Despite the significant impact of pain on population health outcomes, pain remains inadequately assessed in the health care setting [8]. Pain is a complex sensory and emotional experience that is often difficult to communicate [9]. Unidimensional pain measures, such as the numeric or visual analog pain scale, reduce the complex, multifaceted nature of the pain experience to a single number between 0 and 10 [10]. This oversimplification not only results in poor assessment of potential physiological mechanisms but also ignores the complex roles the patient's thoughts and mood play in the patient's pain experience [8].

In some subspecialty medical clinics, multidimensional measures are used, such as the McGill Pain Questionnaire [11], that attempt to take into account other facets of pain beyond intensity, such as pain location, quality, and affective response. However, these measures are often overly complex and rely on long lists of adjectives or esoteric phrases to describe pain that may alienate individuals with low literacy, individuals with dementia or other cognitive limitations, non-native English speakers, and many others with communication limitations [12,13]. With the current state of clinical pain assessment, even individuals without language limitations can have their needs misinterpreted, their symptoms ignored, or their credibility challenged [14]. Ineffective communication about pain may result in patient–clinician discordance, leading clinicians to intervene on poorly described and ill-defined targets, and patients to feel misunderstood and lose trust in their provider [15,16]. The inadequacy of pain assessment tools compromises medical providers' ability to deliver quality care and improve clinical outcomes for their patients [10,17,18].

Painimation

To address the limitations of standard pain assessment, we used human-centered design methods to discover, design, and develop a novel method for assessing pain that leverages digital animations that we call *painimations* [19]. In this work, we hypothesized that an animation-based pain assessment tool would be more acceptable to patients with pain than traditional numerical and adjective-based pain assessments. Our work is particularly timely, given the recent promising evidence suggesting that digital health interventions are feasible, acceptable, and efficacious in a range of chronic medical conditions [20-25].

Our prior publication presented data comparing participants' selection of painimations with their scores on validated, traditional pain scales that rely on pain adjectives and numerical scales [17]. This paper describes our process of using human-centered design to understand how patients experience and express their pain, how clinicians assess and diagnose pain, and how leveraging these observations led to the creation of a novel method for pain assessment: *Painimation*.

Our approach incorporated human-centered design principles, qualitative methods, and stakeholder engagement, and consisted of 3 distinct phases: the inspiration phase, the ideation phase, and the implementation phase [26,27]. After detailing the discovery and development process for a novel, animation-based pain assessment approach, we present initial user testing of the painimations, or abstract animations that can be visually configured to reflect pain quality, pattern, and intensity, as well as the overall Painimation prototype. Finally, we describe future directions for the use of Painimation and discuss how this digital animation approach has the potential to significantly improve medical assessment and treatment of acute and chronic pain.

Methods

Setting

The human-centered design process that resulted in the development of a Painimation prototype took place from January 2015 to May 2016. Key stakeholders were recruited from the Pittsburgh, Pennsylvania, metropolitan area and included patients with acute and chronic pain, clinicians, clinical researchers, and design students. All participants were 18 years of age or older. This project was approved by the University of Pittsburgh's and Carnegie Mellon University's Institutional Review Boards.

Phase I: Inspiration (Empathize, Understand, and Define)

Overview

Human-centered design is inherently an empathic process that attempts to set aside the investigators' or designers' assumptions about the world and gain insight into their users' lived experience, perspectives, pain points, and needs [26]. The goal of Phase I was to *empathize* with the target user and *understand* how pain is experienced and communicated. The next step was to *define* the most prominent barriers to effective patient–provider communication, assessment, and treatment of pain in the health care setting. To accomplish this, we conducted one-on-one, in-depth, in-person interviews with patients with acute and chronic pain, clinicians, and researchers (Table 1).

Table 1. Questions from interviews using directed storytelling and modified think-aloud protocol.

User and stakeholder	Clinician and clinical researcher
A *successful* experience I've had with a clinician around my pain assessment and management was SHORT STORY	The pain assessment protocol I follow is BRIEF OVERVIEW
I describe the pain communication between myself and my clinician as ADJECTIVE	Pain assessment is part of every interaction I have with a patient YES/NO
I summarize my clinician's understanding of and assessment of my pain as ADJECTIVE	I use the following tools LIST/DESCRIBE
I describe my communication ability as ADJECTIVE	I document in the following way ADJECTIVE
I have been asked to rate my pain intensity on a scale like this YES/NO	(Numeric) pain scales are an effective/ineffective CHOOSE tool because REASON
The experience of using the scale was ADJECTIVE	
During that interaction, I communicated the pain intensity that I felt YES/NO	

User and Stakeholder Interviews

Interviews were conducted using directed storytelling [28], a design ethnography method, which allowed patients with a wide range of pain experiences to be interviewed, and yielded information about the contexts in which they had experienced pain as well as descriptions of successful and unsuccessful interactions with their clinicians.

The next part of the patient interviews consisted of a modified version of the think-aloud protocol [29], a method during which participants verbalize their thought process while doing specific tasks. The aim of this portion of the interview was to understand how patients think through 2 current pain scales: the Wong–Baker faces scale and the Numeric Rating Scale [18,30]. Additionally, patients were given a recall interview prompt to understand how they have used these scales in the past to describe their pain to medical providers.

Clinician and Clinical Researchers Interviews

As with the patient interviews, clinician interviews were conducted using directed storytelling to learn about their expertise and experiences in interacting with and treating patients with pain. All interviews were transcribed for later analysis.

Phase II: Ideation (Generate Concepts and Designs)

Overview

The goal of Phase II was to develop solutions to the problem defined in Phase I: how to best allow patients to express their pain and facilitate pain communication with health providers.

Ideation and Concept Development

Analysis of the interviews from Phase I combined thematic analysis and the constant comparison method [31,32]. Codes were developed via open coding of the transcripts to determine topics and themes that emerged. Input from the designers/investigators on relevant topics was also integrated, resulting in a simultaneously inductive and deductive analysis process. Based on the topics identified in Phase I, we developed a set of criteria that needed to be met for a pain communication solution to be considered successful. These criteria served as *design principles* that guided the ideation stage where the designers generated a large number of *concepts*, or creative and innovative solutions to the pain communication problem.

Once several solutions, or concepts, are developed in an unrestricted brainstorm, all of the concepts are evaluated based on the design principles defined earlier. Any concepts that do not meet all of the design principles are discarded. The remaining concepts are ranked relative to 2 axes or factors: importance (ie, potential to impact the problem) and then difficulty (eg, cost, feasibility, scalability). Final concepts are selected based on their relative importance/difficulty and developed using generative storyboards to illustrate how the concepts might function in various scenarios. To test each concept, we conducted needs validation sessions, a design method for working with stakeholders to validate or disprove early ideas, to select a viable concept, and to transition it to the user evaluation stage.

Painimation Drawing Exercises

The process of developing the painimations began with the words used to describe the qualities of pain on the McGill Pain Questionnaire Short Form [11], a pain assessment method that measures pain intensity and quality using 15 descriptors of pain. Drawing exercises were conducted with a group of 16 design students from Carnegie Mellon, to develop visual depictions of the more commonly used pain adjectives. For this exercise, the design students were given a list of qualitative words that are currently used on the McGill Pain Questionnaire Short Form, such as stabbing, pounding, and shooting, and were asked to draw those words, creating a low or medium version, and a high version for each word. The selected words were those most frequently presented by patients and clinicians in Phase I exercises.

Painimation Development

Words from the McGill Pain Questionnaire Short Form [11] were clustered into a few groups, with the idea of creating painimations that would depict and represent different sensations. The first 3 types we explored were throbbing, shooting, and cramping. Deep and dull are terms that could be applied to other qualities, so these were clustered separately. Next, the visual variables that the painimations would represent

or communicate were listed. The final list included speed, saturation, focus, and size (Figure 1). Changing these variables would change the intensity of the pain depicted. These painimations were sent out to the design students in a survey with the question, "What kinds of pain do you believe these animations evoke?" The goal was to understand how participants would describe the qualities of these painimations, given the context of pain.

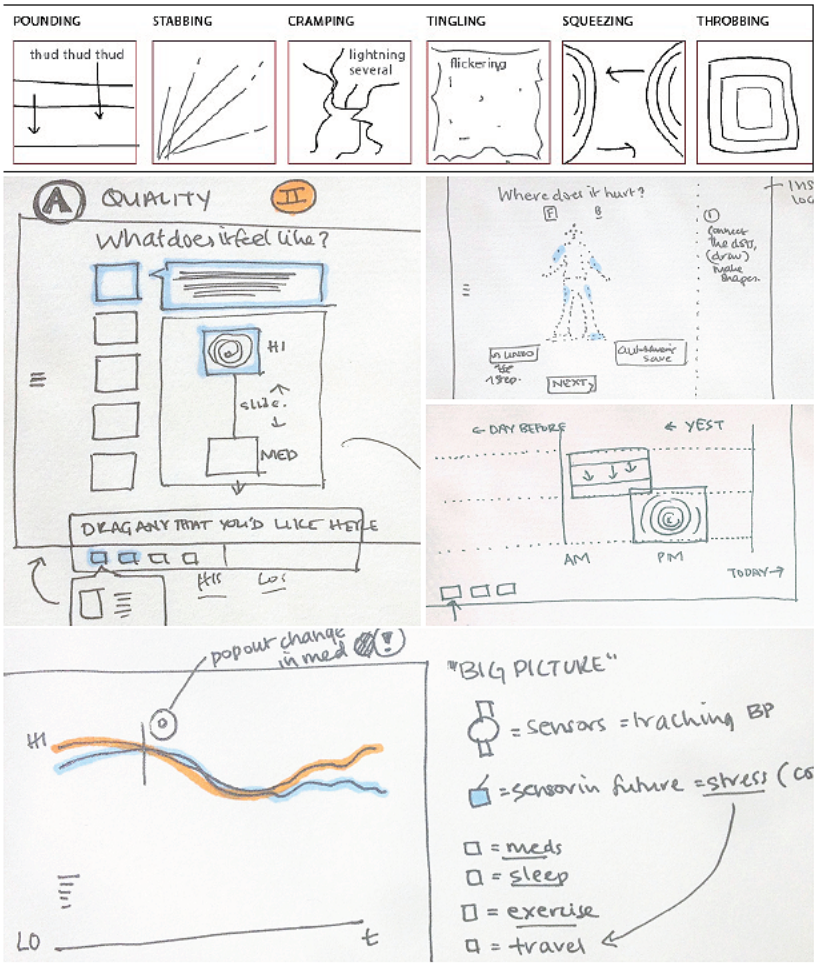
Figure 1. Visual considerations for painimations.

	THROBBING	SHOOTING	CRAMPING
Shape	Squares to indicate sharpness	Lines (pins)	Mimicking electricity
Focus	In focus to communicate intensity	In focus to communicate intensity	In focus to communicate intensity
Color and saturation	Shades of red	Red and black for contrast	White instead of black (to look more like electricity)
Movement and speed	Coming at you to suggest it originates at a certain point and radiates from there	Having a central point but shooting in all directions	Having a central point but shooting in all directions
Fluctuating or consistent	Fluctuating to indicate change over time	Fluctuating to indicate change over time	Fluctuating to indicate change over time
Size	Goes out of frame in certain places, to communicate intensity	Goes out of frame in certain places, to communicate intensity	Stays mostly within the frame

Wireframe Creation

A new set of 11 participants—patients with a history of pain, clinicians, and researchers—were recruited by word of mouth and asked to evaluate the painimations as well as the context of use through the think-aloud protocol. Basic wireframes for the pain assessment app were created to provide context for the painimations (Figure 2). The participants were asked, "How effective do you think this tool is in aiding your pain communication?" via a modified version of the think-aloud protocol. Similarly, clinicians and clinical researchers were asked "Would something like this work? Why or why not?"

Figure 2. Painimation wireframe early sketches.



Phase III: Implementation (Prototype, Test, and Iterate)

Overview

The goal of Phase III was to develop a minimally viable product to test with a small number of users. Once the painimations were refined based on user input from Phase II, we used an agile development process to build a fully functioning *prototype* of an app that utilized the painimations (ie, Painimation). A final set of 8 painimations was developed and subjected to testing and further design iteration. The designers labeled the painimations based on what pain adjective the painimations were intended to represent. Two independent, graduate-level design students were asked to identify what pain type each painimation represented. Confirmation that the painimations approximated the pain adjective they were meant to represent would allow us to transition to pilot testing; otherwise, the painimations would go through another design iteration.

Pilot Testing Using a Case Patient Population: Adults With Sickle Cell Disease

The use-case scenario for Painimation was the assessment and treatment of sickle cell disease (SCD) pain. SCD is a genetic blood disorder that is characterized by unpredictable vaso-occlusive episodes that lead to severe acute pain often called "crisis" and can result in long-term organ damage, chronic pain, and other complications [33]. Patients living with SCD experience pain crisis as early as infancy, and the pain can transition to chronic pain during adolescence and young adulthood. Further, SCD primarily affects underserved, racial/ethnic minorities, and patients often experience discrimination in the medical system [34]. Thus, adults with SCD have long, many times difficult, historical experience with pain and communicating pain to medical providers; these conditions informed the development of this tool.

Participating adult patients with SCD and self-reported chronic pain were presented the 8 painimations and asked, "Would you find this animation applicable to your pain?" These patients were also asked about what types of pain they experience, how they track pain, and their history of pain communication interactions with providers.

Results

Phase I: Inspiration (Empathize, Understand, and Define)

User and Stakeholder Interviews

In total, 10 patients were interviewed, 6 with acute pain (mean age 42.5 years; range 25-50; 50% [n=3] female) and 4 with chronic pain (mean age 40.0 years; range 24-58; 75% [n=3] female). Participants with acute pain had experience with temporary bouts of pain lasting no more than a few days, and patients with chronic pain had a range of pain experiences all lasting more than 3 months. Participants with acute pain experienced a hairline fracture, kidney stones, a pulmonary embolism, postsurgery pain, a root canal, and a urinary tract infection, whereas those with chronic pain experienced migraines, fibromyalgia, vulvodynia, and chronic back pain. Patients with acute and chronic pain both reported having experience communicating pain with clinicians in the medical setting.

Directed storytelling interviews revealed that patients with acute and chronic pain both felt their exact pain was impossible to communicate due to its subjective nature and the individual response to it, both physical and mental. Patients with chronic pain expressed that they particularly struggled to find clinicians who knew and accepted their conditions.

Patients described communication about pain with their health provider as "successful" if they felt heard and understood. Likewise, pain communications were described as "unsuccessful" if there was a lack of understanding, feelings of being dismissed, or intimidated. Textbox 1 displays extracted quotes from these interviews.

The think-aloud protocol revealed that patients with chronic and acute pain both expressed some confusion around traditional pain scales because they felt these scales were "vague" and "ambiguous." For example, several patients stated they had "no clue" what "worst possible pain" in the numerical pain scale meant.

Additionally, patients felt that traditional pain scales "lack specificity" and do not accommodate detailed answers. For example, on the numerical pain scale, one might want to say, "It's an 8 when I am applying pressure, and a 7 when I am resting, and a 10 early in the morning." Patients said that they used these scales to communicate their pain intensity because they had to; 5/10 respondents said their numerical pain rating did not feel accurate.

Textbox 1. Extracted descriptors for clinical communication.

Successful

- Personable, friendly
- Professional
- Dead-on
- Light at the end of a tunnel
- Calming
- Relieving
- I felt in control
- I was actually being heard
- Improved over time

Unsuccessful

- Zero understanding
- Accused me of lying
- Impossible
- Dismissive
- Limited
- Intimidating
- I felt stupid

Clinician and Clinical Researchers Interviews

A total of 7 individuals were interviewed, 4 clinicians (mean age 36.3 years; range 30-50 years, 50% [n=2] female) and 3 clinical researchers (mean age 50.7 years; range 36-58 years; 33% [n=1] female). Clinicians had experience in emergency medicine, general medicine, and physical therapy while clinical researchers had experience in clinical psychology, hematology, and anesthesiology. Clinicians had experience caring for patients with chronic and acute pain, while clinical researchers provided their clinical experience as well as a rich perspective into current research, challenges, and opportunities.

Directed storytelling interviews revealed clinicians' and clinical researchers' perspectives on traditional pain scales. Clinicians explained that a numeric value on the Numerical Pain Rating Scale is only meant to represent one person's pain: "one person's 5 can be compared to their 9, but you cannot compare two individuals' 9's."

A numeric value is useful for communications between clinicians and provides a system that is well understood universally across the medical system. Numeric scales are especially useful in the context of postsurgery pain when clinicians are not as interested in the number itself as in whether the medication or treatment has been effective in reducing pain. In fact, the numeric scale was designed to provide a system for clinicians to note progression in acute and curable pain. Still, some clinician participants stated that in the emergency room, there is some aversion to the numeric system, because patients may exaggerate or falsify their pain score to receive treatment. There was a general belief from respondents that the emergency room sustains the problem of addiction because they cannot deny opioid treatment to patients who report high pain scores, especially if they have an outpatient opioid prescription.

Phase II: Ideation (Generate Concepts and Designs)

Painimation Concept Development

Based on the thematic insights taken from analysis of the user and stakeholder interviews, we established a set of design principles as criteria to support the creation of concept storyboards. A successful solution to the pain communication problem would meet all criteria listed in Textbox 2.

Textbox 2. Design principles as criteria.

Aid patient in describing pain

Given the scope of this project, the attempt was not to remove patient description or report altogether (with automated pain detection, for example) but rather to support that verbal description.

Quantitative representation of pain

Patients want to know that their qualitative experiences matter as much as the quantitative selection. Clinicians, by contrast, required a number of some type that can indicate pain severity and show treatment-related improvements.

Personalized

Patients need to feel that assessment is personalized to them and their pain thresholds. With chronic pain it is all the more important to allow conversations to address the patient's individual journey and take into account changes in their pain experience over time or even moment to moment.

Concise

Because time is limited (and pain assessment is just one part of the interaction between the patient and clinician), the procedure needs to be short and simple to complete, yet provide the necessary data to guide diagnosis and treatment.

Facilitate the conversation

Based on the study findings, the most prominent stakeholder need was for a tool that would improve the patient–provider interaction by making the communication surrounding pain symptoms easier, and helping patients feel heard and understood. The relationship between the patient and clinician was viewed as the most important aspect of the medical encounter.

From the ideation session and concept selection process as described in the "Methods" section, 3 final concepts were selected and then developed out using generative storyboards to illustrate how the principles might fit into various scenarios. The 3 concepts were (1) expressive pain painimations, where patients would use animations to describe their pain to providers; (2) a personalized pain threshold scale, where rather than being restricted to a 0-10 scale patients would use an app to set their highest and lowest pain based on their own descriptors, words, or numbers; and (3) communication-style matching, where patients would be matched with a provider that fits their communication style. Needs validation sessions revealed that only the painimation concept qualified as both desirable and feasible for both patients and clinicians. Patients felt the painimations were more expressive than words or images alone, had an emotional quality, and even incorporated the fluctuations of pain over time. Clinicians and clinical researchers believed the painimation concept could work in their clinic and felt the concept would help create rapport between patients and clinicians.

In terms of feasibility, patients felt there might be individuals who prefer words over images or may not understand the painimations. It was evident that any tool would need to be very easy to understand. Clinicians and clinical researchers expressed that they would still need a number and method to translate the painimations into a score that can indicate severity or be used to compare with the traditional 0-10 numeric pain scale, or other pain assessment measures.

Painimation Drawing Exercises

As part of the painimation development process, a total of 16 graduate design students participated in a drawing exercise where they were asked to draw a series of pain adjectives from the McGill Pain Scale. The student drawings were then clustered based on approach and also arranged according to intensity. This exercise resulted in drawings that were quite similar. The drawings were grouped by similarity, and the final groupings were used to inform the initial set of painimations (Figure 3).

Figure 3. Clustering of participant drawings.

First, a low-medium intensity version was created for the initial 3 painimations (throbbing, cramping, and shooting). Next, 2 new painimations (pounding and tingling) were created, each with a high and medium value. These 2 words came from the original list and were created to provide more variety in the painimations to allow for a range of responses.

Painimation Development

In addition to the drawing exercise session, the graduate design students participated in a survey to evaluate a preliminary set of painimations based on early findings. The "what kinds of pain do you believe these animations evoke?" survey of throbbing, cramping, and shooting painimations revealed rich language within the responses, which were organized into emergent themes: recall, time + change, and representation (Textbox 3). Participants used the painimations as a starting point to recall pain incidents and memories. They mentioned the temporal or changing nature of pain. Additionally, participants indicated satisfaction and comfort using these painimations to represent a sensation.

The throbbing painimation had the highest responses of 1 particular word, which was "throbbing" (n=11). For the shooting painimation, "quick" and "sharp" had the same number of occurrences (n=5). The cramping painimation had a tie between "dull," "deep," and "slow" (n=2). Because of this lack of convergence, the cramping painimation was revised. Multimedia Appendix 1 displays word frequency in responses.

Textbox 3. Emergent themes from survey of painimations.

RECALL

Participants used the animations as a starting point to recall certain pain incidents and memories.

Example quotes:

Reminds me of when I was having my broken arm bent by a pair of nurses to be put into a cast.

Like when I come in from outside when it is cold and my ears heat up uncomfortably, or if I jam my finger and it swells to the point I can feel my heartbeat in my finger.

TIME + CHANGE

Participants mentioned the temporal or changing nature of pain.

Example quotes:

Pain that fluctuates in intensity.

Very erratic pulsing.

Something that starts out in one area and spreads across the body.

Coming up and then dying back down.

Slowly beginning with mild intensity, rising in a crescendo to a near-blinding, wince-inducing pain.

REPRESENTATION

Participants indicated satisfaction and comfort with using these animations to represent a sensation.

Example quotes:

This feels like it could describe that pain well.

I think the strong visuals might really speak to some people.

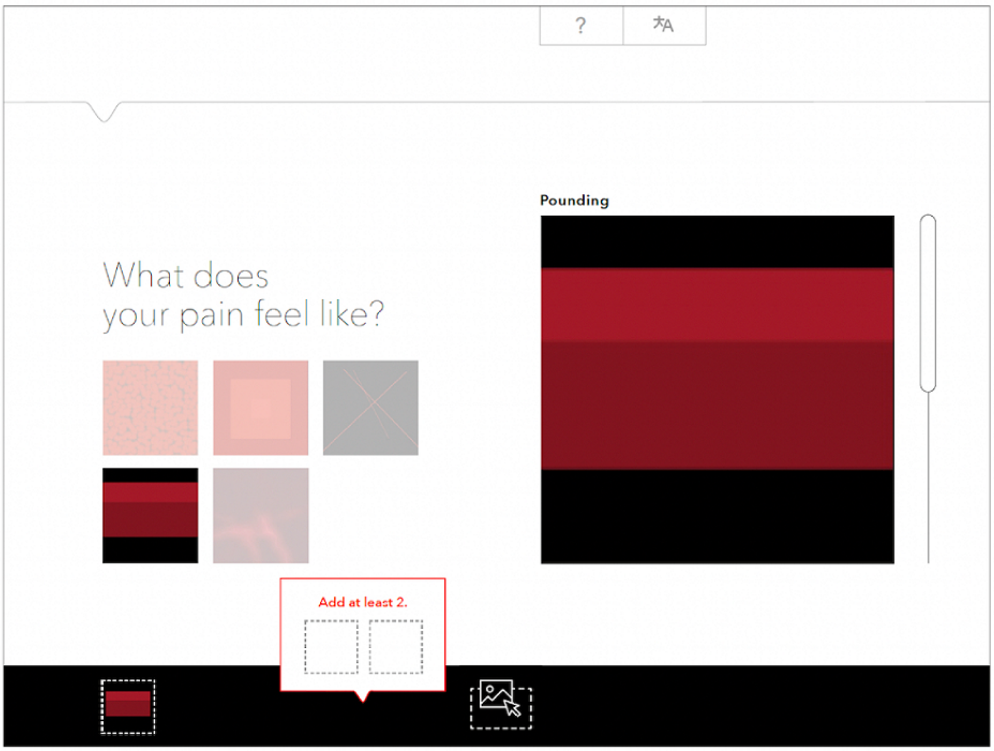
This could easily resemble how I felt when I got my wisdom teeth out.

Creation and Evaluation of Wireframes for a Painimation app

A new group of 5 patient participants and 4 clinician researchers was asked to interact with wireframes of an app that used painimations to measure pain. Feedback included participants wanting to see the whole set of painimations, so they knew how many choices they had. They also preferred that the intensity be depicted through a slider.

To resolve participant concerns, we created an instruction page to precede the viewing of actual images, on how to choose the painimations and increase and decrease the intensity; thumbnails of all painimations were shown on each screen with textual description, and arrows were replaced with a prominent slider. To provide users with feedback after making their selections, a panel was added at the bottom where the chosen painimations could be dragged and dropped (Figure 4).

Figure 4. Painimation selection wireframe.



Phase III: Implementation (Prototype, Test, and Iterate)

Final Set of Painimations

Based on user feedback throughout Phase II, a final set of 8 painimations were developed and then independently reviewed by 2 graduate design students outside of the investigative team (see example 2 in Multimedia Appendix 2).

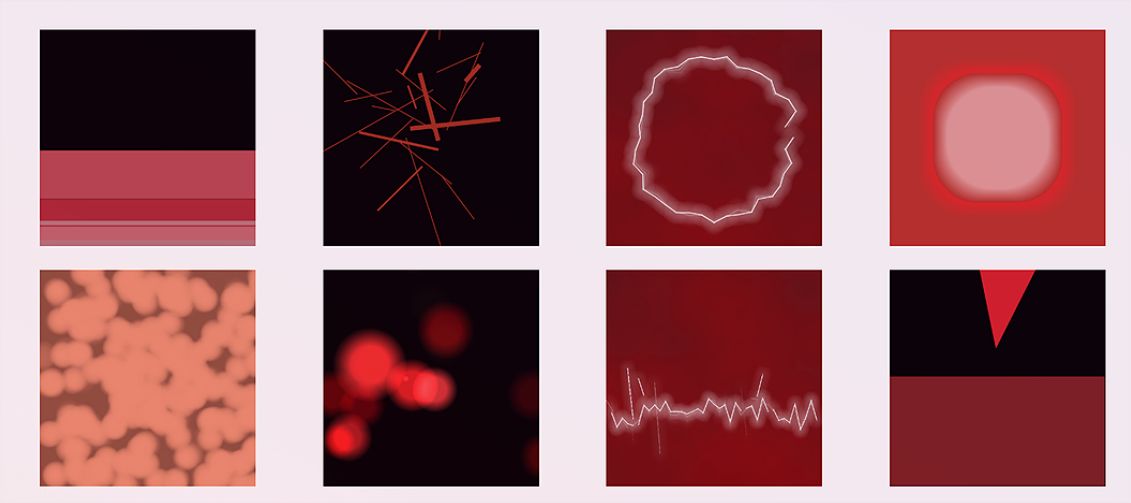
The 2 graduate students were asked to label each painimation using a provided set of pain adjectives. Their labeling of the painimations approximated the intended representations, confirming that a broad set of pain types was depicted as unique feature sets, with no overlap between them.

These final 8 painimations (Figure 5) were then reviewed by the patients (n=5) and clinical researchers (n=4). Patient participants felt these painimations would aid in their pain communication, and several statements suggested that the painimations resonated. Participants would look through the set of painimations, choose 1 or 2, and make statements such as "This one really feels like my headache, exactly!" Other general comments about the idea itself included "These painimations feel like the aha moment for me. Hopefully, doctors will see it soon, too," and "Just knowing that doctors are asking us this question with a tool that comes closer to what we're feeling, shows that they are being empathetic and less dismissive."

While patients acknowledged the benefits of seeing something more qualitative and contextual, they were also concerned about the limitations of the current system: "What is to stop me from getting frustrated with this system in the same way that I currently get frustrated with the number system [wanting to increase the value of the slider to more than what is possible]?"

The clinical researchers' main concern was that the painimations needed validated numerical values of intensity. Although each painimation entry produced a numerical value of 0-100 on the slider and the painimation quality type (eg, "throbbing"), these values would need to have reliable and credible numerical correlations with the traditional numerical pain scale (eg, a particular painimation calibrated at a certain level would equal an 8/10 on the numerical pain scale).

Figure 5. Final set of painimations.



Pilot Testing Painimation With Adults With Sickle Cell Disease

To confirm the acceptability and usability of the painimations, we tested a prototype of a Painimation app designed in Phase II with a use-case sample (adults with SCD-related pain). Six African American adults (age range 24-32 years, 67% [n=4] female) with SCD and self-reported chronic pain completed a pain entry using the prototype Painimation app and were asked to provide a verbal evaluation of all 8 painimations in a modified think-aloud protocol.

The adults with SCD reported that the Painimation app is more engaging, easier to use, has less entry burden, and leads to more of a conversation compared with other pain assessment forms they have used in the past.

In response to the question, "Would you find this animation applicable to your pain?" 6/6 patients with SCD responded "Yes" for electrifying; 5/6 for stabbing; 4/6 for burning; 3/6 for cramping; 2/6 for shooting; and 1/6 for throbbing, tingling, and pounding. Interestingly, 1 participant mentioned that the burning painimation looks like beginning stages of sickle cell crisis. Another patient felt that most painimations were not "severe" enough to represent her pain.

The types of pain seemed to differ between patients; however, many of the patients described their pain as stabbing and pulsating, and they consistently described some of their pain as continuous. In terms of pain tracking, 3/6 patients tracked their pain in their phone or journal, while 2 only documented pain crises, rather than daily pain. One patient said his pain did not change, so he did not feel the need to track it.

These patients echoed what patients with chronic pain in our earlier interviews reported regarding communicating with providers about pain. They liked when they felt like doctors listened and cared but were discouraged when they did not feel heard, when doctors seemed as if they did not have empathy, or did not understand their condition.

Discussion

Application of Painimation

Successful medical care depends on effective communication between patients and clinicians regarding the patients' health symptoms and the most appropriate therapeutic path [35]. Providers are unable to deliver quality medical care when they lack the tools to appropriately assess or interpret patient symptoms that are critical to diagnosis and treatment. This is especially true for the assessment and treatment of pain.

Through a human-centered design approach, our study discovered that patients with pain frequently have negative interactions with providers characterized by misunderstandings, negative accusations, and intimidation. A major cause of this breakdown in the patient–provider interaction is the challenge in communicating pain and feeling understood. Patients, clinicians, and researchers in this study reported that the current pain assessment approaches used in the medical setting fail to accurately capture or communicate patients' pain experience, have limited effectiveness for guiding diagnosis and treatment, and may exacerbate breakdowns in communications between patients and providers. Other studies have also reported that measures oversimplifying the pain experience may lead to patients' personal legitimacy being undermined and result in clinicians inadvertently contributing to chronic pain stigmatization [36]. Given the importance of patients feeling respected and supported by their clinicians, it is imperative to improve patient–clinician communication regarding pain [37].

To address this gap, the current human-centered design study resulted in the development of a novel pain assessment approach that leverages digital animations. The use of pain animations or painimations showed promise with a use-case clinical sample of adults living with SCD-related chronic pain. Our prior published study found that patients' selection of painimations were correlated with their scores on validated scales, and yielded some evidence that painimations may have better diagnostic potential than traditional multidimensional pain scales [17].

Given that pain is incredibly complex and its qualities are particularly difficult to express [8], there have been efforts to improve communication of pain [38-41]. For example, presenting abstract or literal pain images to patients with chronic pain during pain consultations was associated with clinician warmth and empathy, improving the patient–clinician rapport and communication [38]. Another example is Pain QuILT (a newer version of the Iconic Pain Assessment Tool), a web-based and mobile-accessible tool for the visual self-report and tracking of pain that offers 16 pain qualities, such as burning, electrical, and stabbing [39,40]. Pain QuILT was rated significantly easier to use than both the McGill Pain Questionnaire and the Brief Pain Inventory and was associated with fewer barriers to complete [40]. Our findings support and extend this work.

Abstract painimations can capture the experience of pain in a comprehensive manner. These painimations can be visually configured to reflect pain location, quality, and intensity. Moreover, they allow users to interpret the painimations instead of restricting them to specific/labeled pain quality options. The abstract and nonverbal nature of the painimations is also important because it helps level the playing field for marginalized or underserved populations. Patients with lower health literacy, communication disorders, or cultures/languages different from those of the providers have previously faced a communication gap that put them at a disadvantage when seeking medical care. While there are complicated power dynamics between a patient and a clinician, it benefits the patient to have a tool that does not rely on literacy or language, upon which to build conversation and allow patients to more effectively report their symptoms. As evidenced by this project, providing something that is removed from medical jargon or systems (which were not designed from a patient-centric perspective) allows patients to express themselves comfortably, knowing that their comments are valued, heard, and hopefully understood. Furthermore, these painimations address the disparities that current pain assessments perpetuate due to their use of complex words that may alienate individuals with low literacy, disabilities, cognitive impairment, or other communication barriers [12,13,42].

Relevance and Importance of Human-Centered Design Work

Human-centered design and evidence-based data, together, have significant potential for disease prevention and management [43]. Patients need to have the opportunity to participate as true partners in their health care [44]. Utilizing user-centered participatory approaches allows the evaluation of which elements work best for which populations in which contexts [45]. Thus, application of human-centered design in health care will exponentially improve the effectiveness of medical care and disease prevention [43].

Human-centered design is gaining traction in health care and the proliferation of mobile technologies expands opportunities for innovation, particularly because of the wide access to smartphones in clinical populations [23,46-48]. Mobile technologies have been shown to be beneficial in reducing pain severity and are well liked by patients and clinicians [49]. In fact, a study of perspectives of patients with chronic pain on methods of assessing pain found that 80% favored use of a digital version of body template/diagram, and 43% favored use of technology [50]. However, most mobile pain technologies (around 70%) still do not systematically engage patients with chronic pain as end users during app development, nor do they involve clinicians [51]. To ensure short- and long-term engagement of mobile app or digital health interventions, it is critical to include patients and clinicians in all stages, particularly the development stages [48,52-54].

Strengths and Limitations

This study has several strengths, including a rigorous human-centered design approach that involves target users and stakeholders at each phase. A major limitation of this study approach, however, is the small sample and thus limited age, genders, ethnicity/race, and number of pain conditions that were represented by the user and stakeholder groups. For example, only a small number of African Americans with SCD tested the app. Consequently, the generalizability of our findings is limited. The Painimation concept will need to be tested by a larger, more representative sample in terms of age, gender, and ethnicity/race with a broad range of pain types, to determine if all pain experiences are represented in the current set of 8 painimations or if additional painimations need to be designed.

Finally, the reflexivity of the investigators and consultants needs to be considered and was systematically evaluated. It is likely that prior experiences and biases may have influenced the direction of designs and how the findings were interpreted. Future work in this area will benefit from more objective evaluations of the tool and the results.

Future Directions

This study demonstrates the process of human-centered design to build empathy for the end user and ultimately develop and implement an innovative solution for a prominent problem in medical care. Further research is needed to establish whether developing animations that explicitly measure affect and emotion would be beneficial. Additionally, how particular pain characteristics (conditions) might influence the further development of alternative methods (including this one) needs to be considered.

While these painimations have proven to have resonance with participants in this study, there is potential with augmented and virtual reality to develop the pain assessment experience further. For example, a doctoral project at the Norwegian University of Science and Technology in Trondheim is exploring how virtual reality can help nurses develop and sustain their empathy, as clinicians may become desensitized. It simulates morning sickness (nausea and dizziness, for example) through a headset that nurses wear. In relation to this project, the use of painimations in the virtual reality space, recreating the nausea or disorientation that patients with pain would experience, could lead to an intervention that would increase the empathy of family members, friends, and providers towards pain patients.

Finally, with the current data, it is unclear whether Painimation is a tool to replace other measures or to be used in conjunction with other forms of pain assessment. Further, in clinical medicine the 0-10 scale is well-established as the status quo,

and health professionals will need to be convinced that using painimations offers useful and relevant information that can improve their clinical practice. Changing the pain assessment landscape is challenging and there are significant barriers to implementing new tools into routine clinical care. The current body of studies does not address how the pain conversation can be changed in this radical new direction; however, this is a starting point with potential to encourage and inspire other pain researchers to explore novel methods for assessing pain.

Conclusions

This study provides evidence that employing a human-centered design approach in clinical research has the potential to change how medical care is practiced. Currently, most electronic patient-reported outcomes measures for pain are essentially digital copies of paper–pencil questionnaires. Computer adaptive testing has helped streamline assessments, but the fundamental method of assessing symptoms and outcomes with words and numerical scales has not advanced along with the digital era. There is a need for more human-centered design studies to explore how technology can be leveraged to radically improve and advance how patient-reported pain outcomes are assessed.

Acknowledgments

The authors thank Andrea Ball, MA, for her professional medical writing assistance in the preparation of this manuscript.

Conflicts of Interest

CJ is exploring the potential to commercialize Painimation and may open a startup company based on Painimation in the future, but currently receives no revenue or proceeds from Painimation or any products related to this innovation.

Multimedia Appendix 1
Word frequency in responses from the Painimation survey.
[DOCX File , 77 KB - humanfactors_v9i1e27689_app1.docx]

Multimedia Appendix 2
Example of two painimations.
[MOV File , 7070 KB - humanfactors_v9i1e27689_app2.mov]

References

1. Zelaya C, Dahlhamer J, Lucas J, Connor E. Chronic Pain and High-impact Chronic Pain Among U.S. Adults, 2019. NCHS Data Brief 2020 Nov(390):1-8. [Medline: 33151145]
2. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open 2016 Jun 20;6(6):e010364 [FREE Full text] [doi: 10.1136/bmjopen-2015-010364] [Medline: 27324708]
3. Currow D, Agar M, Plummer J, Blyth F, Abernethy A. Chronic pain in South Australia - population levels that interfere extremely with activities of daily living. Aust N Z J Public Health 2010 Jun;34(3):232-239. [doi: 10.1111/j.1753-6405.2010.00519.x] [Medline: 20618262]
4. Schopflocher D, Taenzer P, Jovey R. The prevalence of chronic pain in Canada. Pain Res Manag 2011;16(6):445-450 [FREE Full text] [doi: 10.1155/2011/876306] [Medline: 22184555]
5. Gaskin DJ, Richard P. The economic costs of pain in the United States. J Pain 2012 Aug;13(8):715-724. [doi: 10.1016/j.jpain.2012.03.009] [Medline: 22607834]
6. Kuehn B. Chronic Pain Prevalence. JAMA 2018 Oct 23;320(16):1632. [doi: 10.1001/jama.2018.16009] [Medline: 30357307]
7. Huysmans E, Leemans L, Beckwée D, Nijs J, Ickmans K, Moens M, et al. The Relationship between Cognitive and Emotional Factors and Healthcare and Medication Use in People Experiencing Pain: A Systematic Review. J Clin Med 2020 Aug 03;9(8):2486 [FREE Full text] [doi: 10.3390/jcm9082486] [Medline: 32756298]
8. Craig KD. The social communication model of pain. Canadian Psychology/Psychologie canadienne 2009 Feb;50(1):22-32. [doi: 10.1037/a0014772]
9. Merskey H. Task force on taxonomy. Classification of Chronic: Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms 1986;Suppl 3:226 [FREE Full text]
10. de Williams AC, Davies HTO, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. Pain 2000 Apr;85(3):457-463. [Medline: 10781919]
11. Melzack R. The short-form McGill pain questionnaire. Pain 1987;30(2):191-197. [doi: 10.1016/0304-3959(87)91074-8]
12. Zwakhalen SM, Hamers JP, Abu-Saad HH, Berger MP. Pain in elderly people with severe dementia: a systematic review of behavioural pain assessment tools. BMC Geriatr 2006 Jan 27;6(1):1-15 [FREE Full text] [doi: 10.1186/1471-2318-6-3] [Medline: 16441889]

13. Herr K, Coyne PJ, McCaffery M, Manworren R, Merkel S. Pain Assessment in the Patient Unable to Self-Report: Position Statement with Clinical Practice Recommendations. Pain Management Nursing 2011 Dec;12(4):230-250. [doi: 10.1016/j.pmn.2011.10.002]
14. Hadjistavropoulos T, Craig KD, Fuchs-Lacelle S. Social influences and the communication of pain. In: Hadjistavropoulos T, Craig KD, editors. Pain: Psychological Perspectives. New York, NY: Lawrence Erlbaum Associates Publishers; 2004:87-112.
15. Butow P, Sharpe L. The impact of communication on adherence in pain management. PAIN® 2013;154:S101-S1S7. [doi: 10.1016/j.pain.2013.07.048]
16. Dobkin P, Sita A, Sewitch M. Predictors of adherence to treatment in women with fibromyalgia. The Clinical journal of pain 2006;22(3):286-294. [doi: 10.1097/01.ajp.0000173016.87612.4b]
17. Jonassaint CR, Rao N, Sciuto A, Switzer GE, De Castro L, Kato GJ, et al. Abstract Animations for the Communication and Assessment of Pain in Adults: Cross-Sectional Feasibility Study. J Med Internet Res 2018 Aug 03;20(8):e10056 [FREE Full text] [doi: 10.2196/10056] [Medline: 30076127]
18. Clark WC, Yang JC, Tsui SL, Ng K, Clark SB. Unidimensional pain rating scales: a multidimensional affect and pain survey (MAPS) analysis of what they really measure. Pain 2002 Aug;98(3):241-247. [doi: 10.1016/S0304-3959(01)00474-2] [Medline: 12127025]
19. Rao N. Redesigning the Pain Assessment Conversation (PhD Thesis). 2015. URL: https://s3-eu-west-1.amazonaws.com/pstorage-cmu-348901238291901/12259871/RedesigningthePainAssessmentConversation.pdf [accessed 2021-12-07]
20. Badawy SM, Radovic A. Digital Approaches to Remote Pediatric Health Care Delivery During the COVID-19 Pandemic: Existing Evidence and a Call for Further Research. JMIR Pediatr Parent 2020 Jun 25;3(1):e20049 [FREE Full text] [doi: 10.2196/20049] [Medline: 32540841]
21. Radovic A, Badawy SM. Technology Use for Adolescent Health and Wellness. Pediatrics 2020 May 01;145(Supplement 2):S186-S194. [doi: 10.1542/peds.2019-2056g]
22. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: A systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan 19;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]
23. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-Centered eHealth Interventions for Children, Adolescents, and Adults With Sickle Cell Disease: Systematic Review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
24. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
25. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
26. IDEO. The field guide to human-centered design. New York, NY: Design Kit; 2015:311.
27. Brown T, Wyatt J. Design Thinking for Social Innovation. Development Outreach 2010 Jul;12(1):29-43. [doi: 10.1596/1020-797x_12_1_29]
28. Evenson S. Directed storytelling: Interpreting experience for design. Design Studies: Theory and research in graphic design 2006:231-240.
29. Jääskeläinen R. Think-aloud protocol. Handbook of translation studies 2010;1:371-374. [doi: 10.1075/hts.1.thi1]
30. Garra G, Singer AJ, Domingo A, Thode HC. The Wong-Baker Pain FACES Scale Measures Pain, Not Fear. Pediatric Emergency Care 2013;29(1):17-20. [doi: 10.1097/pec.0b013e31827b2299]
31. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, She KJ, editors. APA Handbook of Research Methods in Psychology: Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington, DC: American Psychological Association; 2012:57-71.
32. Glaser BG. The Constant Comparative Method of Qualitative Analysis. Social Problems 1965 Apr;12(4):436-445. [doi: 10.2307/798843]
33. Ware RE, de Montalembert M, Tshilolo L, Abboud MR. Sickle cell disease. The Lancet 2017 Jul;390(10091):311-323. [doi: 10.1016/s0140-6736(17)30193-9]
34. Fuggle P, Shand PA, Gill LJ, Davies SC. Pain, quality of life, and coping in sickle cell disease. Arch Dis Child 1996 Sep 01;75(3):199-203 [FREE Full text] [doi: 10.1136/adc.75.3.199] [Medline: 8976657]
35. Street RL. How clinician-patient communication contributes to health improvement: modeling pathways from talk to outcome. Patient Educ Couns 2013 Sep;92(3):286-291. [doi: 10.1016/j.pec.2013.05.004] [Medline: 23746769]
36. Cohen M, Quintner J, Buchanan D, Nielsen M, Guy L. Stigmatization of patients with chronic pain: the extinction of empathy. Pain Med 2011 Nov 01;12(11):1637-1643. [doi: 10.1111/j.1526-4637.2011.01264.x] [Medline: 22054062]
37. Henry S, Matthias M. Patient-clinician communication about pain: a conceptual model and narrative review. Pain Medicine 2018;19(11):2154-2165. [doi: 10.1093/pm/pny003]

38. Ashton-James CE, Dekker PH, Addai-Davis J, Chadwick T, Zakrzewska JM, Padfield D, et al. Can images of pain enhance patient–clinician rapport in pain consultations? British Journal of Pain 2017 Jun 26;11(3):144-152. [doi: 10.1177/2049463717717125]
39. Lalloo C, Henry JL. Evaluation of the Iconic Pain Assessment Tool by a Heterogeneous Group of People in Pain. Pain Research and Management 2011;16(1):13-18. [doi: 10.1155/2011/463407]
40. Lalloo C, Kumbhare D, Stinson JN, Henry JL. Pain-QuILT: Clinical Feasibility of a Web-Based Visual Pain Assessment Tool in Adults With Chronic Pain. J Med Internet Res 2014 May 12;16(5):e127. [doi: 10.2196/jmir.3292]
41. McAuley K. Improving communication in the diagnosis of painful neuropathy. J Diabetes Nurs 2009;13(2):46-56 [FREE Full text]
42. Hadjistavropoulos T, Craig K. A theoretical framework for understanding self-report and observational measures of pain: a communications model. Behaviour Research and Therapy 2002 May;40(5):551-570. [doi: 10.1016/s0005-7967(01)00072-9]
43. Matheson GO, Pacione C, Shultz RK, Klügl M. Leveraging Human-Centered Design in Chronic Disease Prevention. American Journal of Preventive Medicine 2015 Apr;48(4):472-479. [doi: 10.1016/j.amepre.2014.10.014]
44. Murphy J. Patient as center of the health care universe: A closer look at patient-centered care. Nursing economics 2011;29(1):35. [Medline: 21469488]
45. Lor M, Koleck T, Bakken S. Information visualizations of symptom information for patients and providers: a systematic review. Journal of the American Medical Informatics Association 2019;26(2):162-171. [doi: 10.1093/jamia/ocy152]
46. Bazzano AN, Martin J, Hicks E, Faughnan M, Murphy L. Human-centred design in global health: A scoping review of applications and contexts. PLoS ONE 2017 Nov 1;12(11):e0186744. [doi: 10.1371/journal.pone.0186744]
47. Holeman I, Kane D. Human-centered design for global health equity. Inf Technol Dev 2019 Sep 29;26(3):477-505 [FREE Full text] [doi: 10.1080/02681102.2019.1667289] [Medline: 32982007]
48. Badawy SM, Thompson AA, Liem RI. Technology Access and Smartphone App Preferences for Medication Adherence in Adolescents and Young Adults With Sickle Cell Disease. Pediatr Blood Cancer 2016 May 04;63(5):848-852. [doi: 10.1002/pbc.25905] [Medline: 26844685]
49. Thurnheer SE, Gravestock I, Pichierri G, Steurer J, Burgstaller JM. Benefits of Mobile Apps in Pain Management: Systematic Review. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e11231 [FREE Full text] [doi: 10.2196/11231] [Medline: 30348633]
50. Verge AC, Mukhida K. Patients' perspectives on methods of assessing pain. DMJ 2020 Jul 19;46(2):e10144. [doi: 10.15273/dmj.vol46no2.10144]
51. Zhao P, Yoo I, Lancey R, Varghese E. Mobile applications for pain management: an app analysis for clinical usage. BMC Med Inform Decis Mak 2019 May 30;19(1):106 [FREE Full text] [doi: 10.1186/s12911-019-0827-7] [Medline: 31146739]
52. Badawy SM, Thompson AA, Kuhns LM. Medication Adherence and Technology-Based Interventions for Adolescents With Chronic Health Conditions: A Few Key Considerations. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e202 [FREE Full text] [doi: 10.2196/mhealth.8310] [Medline: 29273573]
53. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
54. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]

Abbreviations

CDC: Centers for Disease Control and Prevention
SCD: sickle cell disease

Edited by A Kushniruk; submitted 03.02.21; peer-reviewed by S Badawy, Y Virella Pérez, B Thompson, R de la Vega; comments to author 03.03.21; revised version received 17.03.21; accepted 06.09.21; published 07.01.22.

Please cite as:
Rao N, Perdomo S, Jonassaint C
A Novel Method for Digital Pain Assessment Using Abstract Animations: Human-Centered Design Approach
JMIR Hum Factors 2022;9(1):e27689
URL: https://humanfactors.jmir.org/2022/1/e27689
doi:10.2196/27689
PMID:34994697

Original Paper

Web-Based Structured Education for Type 2 Diabetes: Interdisciplinary User-Centered Design Approach

Shoba Poduval[1], MRCGP, PhD; Jamie Ross[1], PhD; Kingshuk Pal[1], MRCGP, PhD; Nikki Newhouse[2], PhD; Fiona Hamilton[1], MRCGP, PhD; Elizabeth Murray[1], FRCGP, PhD

[1]Research Department of Primary Care & Population Health, University College London, London, United Kingdom

[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Shoba Poduval, MRCGP, PhD
Research Department of Primary Care & Population Health
University College London
Royal Free Hospital, Rowland Hill Street
London, NW3 2PF
United Kingdom
Phone: 44 020 3108 1207
Email: s.poduval@ucl.ac.uk

Abstract

Background: Digital health research encompasses methods from human-computer interaction and health research.

Objective: This paper aims to describe how these methods were combined to develop HeLP-Diabetes: Starting Out, a web-based structured education program for people newly diagnosed with type 2 diabetes.

Methods: The development process consisted of three phases: initial design for effectiveness, optimization for usability, and *in the wild* testing in the National Health Service with people newly diagnosed with type 2 diabetes, and further revisions. We adopted an iterative user-centered approach and followed steps from the human-computer interaction design life cycle and the Medical Research Council guidelines on developing and evaluating complex interventions.

Results: The initial design process resulted in an 8-session program containing information and behavior change techniques targeting weight loss, being more active, and taking medication. The usability testing was highlighted at an early stage, where changes needed to be made to the language and layout of the program. The *in the wild* testing provided data on uptake of and barriers to use. The study suggested low uptake and completion of the program, but those who used it seemed to benefit from it. The qualitative findings suggested that barriers to use included an expectation that the program would take too long. This informed refinements to the program.

Conclusions: The use of interdisciplinary methods resulted in an iterative development process and refinements to the program that were based on user needs and data on uptake. The final intervention was more suitable for a definitive evaluation than the initial version. The description of our approach informs other digital health researchers on how to make interventions more sensitive to user needs.

(JMIR Hum Factors 2022;9(1):e31567) doi:10.2196/31567

KEYWORDS

type 2 diabetes; patient self-management; diabetes education; primary care; digital health

Introduction

Interdisciplinary Research Methods

Research on digital health interventions (DHIs) brings together the human-computer interaction (HCI; which includes software engineering) and health (encompassing biomedical, behavioral, and social sciences). The research methods used in HCI, such as health research, are largely empirical (eg, experimental designs, surveys, and focus groups). However, health research tends to use a sequential approach, based on the methods used in pharmacological drug development, culminating in a randomized controlled trial to determine its effectiveness [1]. In HCI research, there is more emphasis on proximal (interaction) and distal (effects) outcomes, and the need to iteratively design and test an intervention until it is deemed to

be *accessible and useful* by the user [2-4]. Acceptability and usability are crucial to digital health researchers, because the effectiveness of DHIs relies on being used (at the individual level), and the population impact depends on reaching a high proportion of the target population. The use of iterative methods common to HCI allows DHIs to be optimized until they are likely to achieve sufficient acceptability to ensure adequate reach, uptake, and use to achieve effectiveness and cost-effectiveness [5]. A decision about whether to proceed to a definitive randomized trial can then be made.

The Medical Research Council (MRC) has published guidelines for health researchers researching complex interventions to help them adopt appropriate methods. The 2006 MRC framework suggested a nonlinear approach to the development and evaluation of complex interventions, with four key stages: (1) development, (2) feasibility and piloting, (3) evaluation, and (4) implementation [6].

These 4 stages involve using evidence and theory to develop complex interventions, then testing them with a series of pilot studies aimed at key design uncertainties, before moving on to an exploratory and then a definitive evaluation [6]. MRC best practice guidelines is that definitive evaluation should only be undertaken once (1) the intervention and its delivery package reach a degree of stability, (2) any further development would be relatively minor, (3) there is reasonable confidence that the intervention could be implemented with high fidelity, and (4) there is a reasonable likelihood that the intervention will lead to improved health outcomes or equivalent outcomes at lower cost [5].

Iterative development and evaluation are also features of the HCI development life cycles. A life cycle is the sequence of activities that occurs from the initial concept, through to the eventual phasing out and replacement [7]. The process purposefully cycles through several designs, incrementally improving the design until the final product is reached [7]. A key aspect is to be user-centered and involve users throughout the design process. This allows designers to understand people in the contexts in which they live, work, and learn, and consequently how to design products that fit easily into users' everyday lives [8]. Standard frameworks for HCI and usability have been developed that recommend an iterative design process with an emphasis on the continuous identification of user requirements, testing the intervention against these requirements, respecifying user requirements, and retesting [9,10]. These processes inform each other and are repeated in each design cycle (Figure 1).

The challenges of interdisciplinary work across HCI and health have been highlighted by Pagliari [11] and Blandford et al [2]. One of the key issues is that, although both the MRC and the HCI life cycle approaches are iterative, the HCI life cycle is located entirely in the development phase of the MRC framework, as illustrated in Figure 2.

Figure 1. The human-computer interaction design cycle [10]. Used with permission from Elsevier.

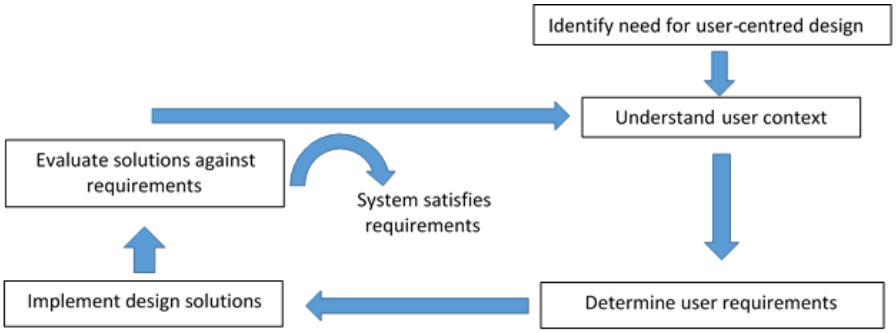
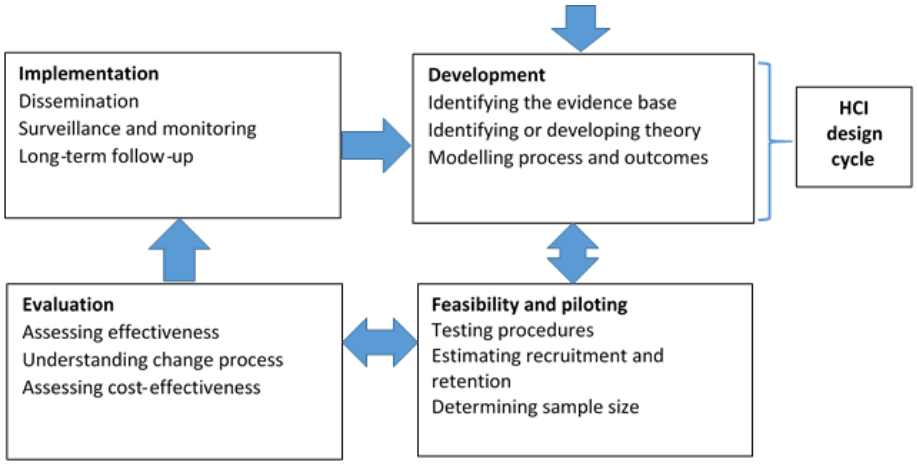


Figure 2. The Medical Research Council framework and human-computer interaction design cycle [1]. Used with permission from the BMJ Publishing Group Ltd. HCI: human-computer interaction.



This paper focuses on this development phase, with the aim of highlighting the importance of careful user-centered design and early-stage usability testing before further evaluation. We have reported how we tackled this, by describing the 3 phases of development and describing the methods from HCI and the methods from health that we incorporated at each phase. The aim of reporting our methods is to guide other researchers in developing similar interventions. Historically, complex interventions, such as diabetes self-management programs, have not been well described [12,13]. Better reporting would improve the understanding of causal mechanisms and increase the collective knowledge of how and why interventions work or not. This would, in turn, facilitate learning among researchers and the development of more effective interventions [14].

Our approach was also informed by guidelines on evaluating DHIs [5], which defines the research questions forming the basis of an evaluation of a DHI, and the issues that are particularly salient to DHIs rather than complex interventions as a whole. The guidelines on DHI evaluation was particularly relevant during the early stages of design when we needed to identify the health needs, target population, and causal model for the intervention.

There are other examples where the process of combining interdisciplinary methods have been documented [15-19]. The person-based approach developed by Yardley et al [15], for example, reports a development process involving qualitative interviews with a wide range of people from the target user population at every stage. Insights from users are then used to modify the intervention to make it more persuasive, feasible, and relevant. In keeping with the person-based approach, we collected insights from users and used these insights to modify the intervention at both the usability testing and *in the wild* testing stages.

Background to the Intervention

Guidelines for evaluating DHIs recommend starting by defining the problem to be addressed, namely the health need that the DHI is intended to address and the population who could benefit from the DHI. For HeLP-Diabetes: Starting Out (HDSO), the health need that is being addressed is the provision of a structured education for people with type 2 diabetes mellitus (T2DM). T2DM is an international priority, affecting approximately 425 million people worldwide. T2DM places a considerable burden on patients in terms of premature morbidity and mortality and on health services, in terms of cost. Both these burdens can be reduced by structured self-management education, which can improve patient knowledge, self-care behaviors, metabolic control, psychological outcomes, and health care costs [20-23]. In the United Kingdom, T2DM affects an estimated 3.8 million people aged >16 years (8.6% of the population of this age group) [24] and accounts for approximately 10% of the total National Health Service (NHS) budget [25].

It is an NHS policy that all patients diagnosed with T2DM are offered structured education [26]. General practitioners (GPs) in England are remunerated through the Quality and Outcomes Framework (QOF) for referring newly diagnosed patients to suitable programs, with the suitability of the program determined by the accreditation by the Quality Institute for Self-Management Education (QISMET) [27,28]. Despite this incentivization, the uptake of structured education is poor (8.3% uptake in 2016 [29]). The reasons for this low uptake include difficulties with the current dominant model of structured education which is group-based and can be difficult for people who work, have caring responsibilities, or dislike groups. Our team had already developed a web-based self-management program (HeLP-Diabetes) [30], which was shown to be effective and cost-effective [31,32]. HeLP-Diabetes is a website with over 560 pages that provide self-management support for patients from diagnosis to death. The content is broken down into 8 sections, including information about understanding and treating diabetes, an interactive health record, news and research, and a forum and help page [31]. Engagement was also encouraged with regular emails and text that contained links to topical content within the website (eg, information regarding influenza vaccinations in winter) [30,33].

However, QOF payment and QISMET accreditation require a structured program (with a clear curriculum and learning goals and modules to work through in a linear fashion) aimed at newly diagnosed patients. HeLP-Diabetes was not structured (people have access to the website, without following a linear pathway), and it was not aimed at newly diagnosed patients but at patients at all stages of their diabetes journey. Therefore, we decided to develop a web-based structured course that could gain QISMET accreditation and meet the QOF requirements. The established courses that the GPs could refer patients to and gain QOF remuneration were all group-based and face-to-face; thus, a web-based structured course would provide an alternative that could potentially bypass some of the barriers to uptake described earlier. Table 1 illustrates the key differences between the HeLP-Diabetes website and the HDSO-structured course described in this paper.

Table 1. Key differences between HeLP-Diabetes and HeLP-Diabetes: Starting Out.

Feature	HeLP-Diabetes	HeLP-Diabetes: Starting Out
Target user	People with T2DM[a] at any stage	Newly diagnosed people with T2DM
Size	8 sections, with 560 pages	5 sections, with selected content from HeLP-Diabetes
How the intervention was delivered	Nonlinear—people could access any part of the website and dip in and out as they pleased	Linear—people worked through modules one by one, and were given access to the next module once they completed the previous one
Curriculum	No curriculum—a wide breadth of information was available, and people could choose which topics to access depending on interest	Spiral curriculum—people worked through a series of modules and added to the knowledge they gained from previous modules in a *spiral* fashion

[a]T2DM: Type 2 diabetes mellitus.

The interdisciplinary development process required the skills of a multidisciplinary team of patients, GPs, diabetes nurse specialists, and health and HCI researchers. KP and JR formed part of the HeLP-Diabetes team, and KP and SP formed part of the HDSO team. EM led both teams.

Aim and Objectives

This paper aims to describe how we combined the methods described in the MRC and HCI guidelines, using the development of the HDSO program as a worked example. The three stages of development we undertook were as follows: (1) phase 1—initial design, (2) phase 2—optimizing for usability, and (3) phase 3—*in the wild* testing and further revisions.

For each stage of development, we have described the methods from HCI, the methods from health, and how we combined the two. The evaluation of the final intervention for feasibility, acceptability, and impact is described elsewhere [34].

Methods

Phase 1: Design for Effectiveness

Methods From HCI: Establishing User Requirements for HeLP-Diabetes as a Precursor for HDSO

Focus Groups

The first steps of the HCI design process involve understanding the user context and requirements. This took place during the development of the HeLP-Diabetes website, before the development of the HDSO-structured program.

Understanding the contexts in which people live, work, and learn allows designers to develop products that fit easily into users' everyday lives. Products that are easy to use are more likely to be acceptable to patients and more widely taken up. Extensive work in establishing the requirements of patients with T2DM went into the development of HeLP-Diabetes, the precursor to HDSO, and has been reported by Dack et al [30]. User requirements were conceptualized as features that would make people want to use the interventions (*wants*) and features needed to help improve health outcomes (*needs*). The HeLP-Diabetes team conducted focus groups with patients and health professionals (health professionals facilitated engagement with the program) to collect this information.

Usability Testing

The content identified as necessary in the focus groups was integrated by the design team and then reviewed by a participatory design group consisting of patients with T2DM. The content went through several iterations and was put through usability testing. Usability testing is commonly used in software engineering and HCI research. It has been described as "representative users attempting representative tasks in representative environments, on early prototypes or working versions of computer interfaces" [35]. Usability testing aims to find flaws in the interface that need improvement and to make products more sensitive to users' needs at an early stage of development [36]. There is a wide range of techniques used in usability testing, including questionnaires, think-aloud observation, and interview-based techniques [37]. Usability testing for HeLP-Diabetes involved users *thinking aloud* while undertaking prespecified tasks (eg, finding specific information or using one of the self-monitoring tools), a technique common and unique to HCI [38]. Usability testing helped to optimize the navigation and interactive features of HeLP-Diabetes. Selected content from HeLP-Diabetes were used to develop the HDSO program, informed by evidence, theory, and modeling.

Methods From Health

Evidence

Systematic reviews of web-based diabetes self-management interventions [39-42] have found that the most effective components are (1) prompting of self-monitoring of behavioral outcomes, (2) provision of information on consequences of behavior, (3) barrier to identification or problem solving, (4) feedback on performance, and (5) interaction with health care professionals via the internet [39,40,42]. This evidence was combined with the theory regarding long-term condition self-management to determine the necessary components of the program. This theory is discussed in the next section.

Theory and Causal Modeling

The aim of the structured program was not only to impart knowledge but also to empower and encourage people newly diagnosed with T2DM to improve their self-efficacy (self-confidence in self-management) and emotional well-being by learning about living a healthy lifestyle, making the most of the NHS and staying motivated. There were many theories and theoretical models which related to the aims of the program. These included the Corbin and Strauss model for the work of

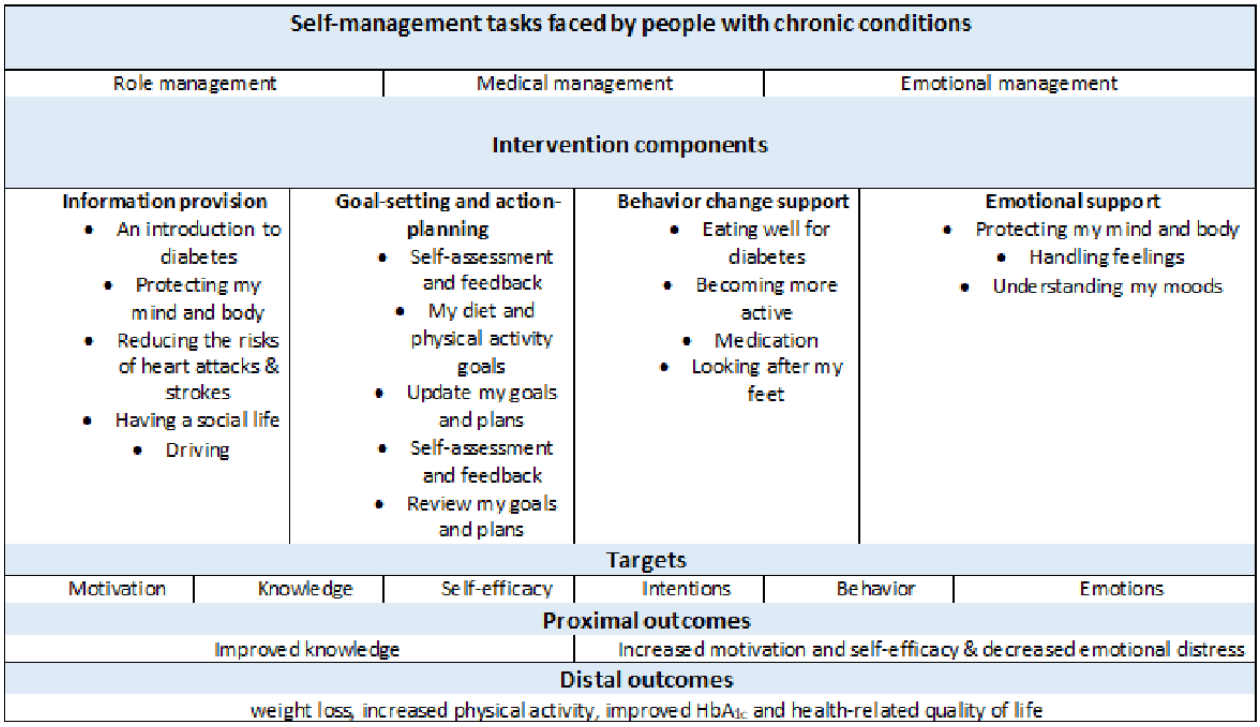
living with a long-term condition [43] and behavior change theories. The Corbin and Strauss model was chosen because of its holistic approach to diabetes self-management and fit with education theory about multidimensional learning. Increasingly, learning has been construed as being multidimensional and involving the body, emotions, spirit, and the mind [44]. The Corbin and Strauss model for the work of living with a long-term condition also emphasizes the need to address the emotional aspects of disease and identity issues. Corbin and Strauss identified 3 sets of tasks involved in self-management [45] from qualitative work on the perception of patients about their long-term conditions. These are conceptualized as follows [43]:

1. Medical management: adopting healthy behaviors (eg, not smoking, exercising regularly, and eating healthy food), working with health professionals (eg, keeping appointments and following instructions), and taking medicines.
2. Emotional management: addressing the negative emotions associated with being diagnosed with a long-term condition.
3. Role management: coming to terms with the disruption to one's sense of self, including adjusting to the *patient* role and managing the impact of one's diagnosis on relationships with friends, family, and colleagues.

Behavior change theories were used because they can help predict how and when behavior change occurs [46]. Behavior change techniques (BCTs) are the strategies used in an intervention to promote behavior change [47]. They can be designed using behavior change theories. Interventions that use more theory-based BCTs have been found to have larger effect sizes compared with interventions that use fewer techniques in studies of digital health behavior change interventions [47].

Guidelines on evaluating DHIs [5] recommends identifying the necessary components of an intervention (including BCTs) by establishing a credible causal pathway for the intervention; thereby, linking evidence and theory to the intended outcomes. We linked the 3 self-management tasks identified in the Corbin and Strauss model to the intended outcomes of the intervention (improved knowledge, self-efficacy, and emotional well-being) using a causal modeling approach. The causal model for HDSO is illustrated in Figure 3.

Figure 3. Causal model of HeLP-Diabetes: Starting Out program. HbA_{1c}: glycated hemoglobin A_{1c}.



In the long term, the proximal outcomes of knowledge, motivation, self-efficacy, and emotional distress combine to enable behavior change and better glycemic control. We opted not to measure long-term outcomes but focus on the short-term outcomes, including registration, use, and change in these 3 proximal outcomes.

How HCI and Health Methods Were Combined

Findings from the work on user requirements (HCI) and evidence, theory, and causal modeling of the intervention (health) were combined to select appropriate BCTs and develop the content, format, and structure of the program.

Phase 2: Optimize for Usability

Methods From HCI

The usability testing of the HDSO program was conducted using questionnaires emailed to 5 patient volunteers. The questionnaires were written by the GPs and diabetes specialist nurses in the HDSO team and included the following four items: (1) the title of the session (eg, easy to understand and relevant to content), (2) the information contained in the session (eg, appropriate quantity of information, encouraging tone, and relevant links to information from other sources), (3) the visual design of the session (eg, readability of font, ease of finding

access to videos, and the next section of the program), and (4) any other specific suggestions.

Methods From Health

Usability testing formed part of the early phase work recommended by the MRC and was undertaken to test the initial design with target users, before the exploratory *in the wild* testing with patients in the NHS.

How HCI and Health Methods Were Combined

Combining the phased approach advocated by the MRC and making changes to the program based on the results of 1 round of usability testing meant that the early development of the program was iterative. This was a key strength of this research. We did not immediately proceed to evaluation after the initial design of the program but undertook cycles of testing, refinement, retesting, and further refinement until we were more confident that the intervention fulfilled user needs and did not need major changes. This was recommended by both the MRC guidelines used by health researchers and the HCI lifecycle models used by HCI researchers. The refinements made to the design of the program were based on the results of the usability testing (and therefore based on user needs and experience).

The purpose of the next stage of testing and refinement was to evaluate the design of the program against user requirements. This was undertaken using *in the wild* testing.

Phase 3: In the Wild Testing and Further Revisions

Methods From HCI

Research *in the wild* is a term used for research conducted in natural settings. It is increasingly used in HCI to understand how people react to and integrate technologies in their everyday lives over a period [48,49]. In situ studies are more likely to reveal the behaviors people adopt and the problems they encounter when they use an intervention at home, at work, or elsewhere. The advantage of this is that they provide greater external validity than experimental studies, where participants are more aware of how they are expected to behave. Another advantage is that *in the wild* studies nearly always provide unexpected findings regarding what humans do when confronted with a digital intervention; these can be the most informative findings [49]. In addition to usability testing, this provided an extra way of testing interventions against user requirements (as suggested in the HCI design cycle).

Methods From Health

The aim of the *in the wild* testing was to understand more about how patients were using the HDSO program in their everyday lives, including their experiences and views of the problems they encountered in using the program. Therefore, mixed methods were used in this study. Quantitative data were collected on the number of patients registering for and completing the program, patient characteristics, and changes in questionnaire scores. The questionnaires administered were the Problem Areas in Diabetes measuring diabetes-related distress [50], and the diabetes management self-efficacy questionnaire [51]. Questionnaires were included in the web-based program at the start and end of the course. Qualitative methods were used to explore the patient experiences and views regarding using the program.

Qualitative methods are used in both health and HCI research, but there are some important differences in the approaches used. For example, the *locus of expertise* differs. In health and social science research, researchers typically start with their own expertise rather than the user's expertise, and the design interventions that (it is hoped that) users will engage with [2]. In HCI research, the user is assumed to be an expert in what they do and what they need. Digital health research has adopted more *user-centered* approaches to address the challenge of low uptake and adherence, with a focus on understanding and accommodating the perspectives of the people who will use the intervention [15]. In health and social science, a less formative (developmental) and more summative (cumulative) approach is often taken, so that there is less focus on early outcomes in the developmental stages and more focus on the impact of the final intervention. Emphasis is placed on conducting interviews of sufficient depth and duration. Interviews in HCI research use methods that are more common to industry and are driven by time and resources. Rapid user experience studies with smaller sample sizes are conducted at several stages during product development.

How HCI and Health Methods Were Combined

Methods from HCI and health were combined to conduct the *in the wild* testing of the HDSO program. The setting for the *in the wild* testing was GP practices. This was the natural setting for this study as referrals to structured education for T2DM patients occurs in primary care. Practices in 2 London boroughs that had taken part in a HeLP-Diabetes implementation study and practices in 2 London boroughs that were interested in commissioning the HDSO program participated in the study. The program was offered to these practices for free as an alternative to established face-to-face diabetes structured education courses that were already commissioned.

The study was submitted to the Health Research Authority (HRA) for NHS Research Ethics Committee (REC) for review. Secondary analysis of information collected as part of normal care was excluded from the REC review by the HRA as long as the patients were not identifiable [52]. Therefore, the collection of data on registrations, completed sessions, and questionnaire scores were permissible, as the data were automatically pseudonymized with a numerical identifier. Patients were informed on registration that anonymized data were collected by the program and used anonymously for ongoing service development.

A total of 15 practices agreed to offer the HDSO program to patients for the study. The program was offered to patients as an NHS service; therefore, there were no formal inclusion and exclusion criteria. Practices were informed that the target population of the intervention was adults (aged ≥18 years) with T2DM diagnosed in the last 9 months and asked to offer the program to everyone in this population. Practices were asked to identify eligible patients by running a search of the electronic medical records. Practices were sent registration packs to mail out to eligible patients. A total of 322 packs were mailed out. The registration pack contained information about the HDSO

program (including that it was being offered as an NHS service as an alternative to face-to-face courses), how to register, and a reply slip. Patients interested in using the program returned a reply slip to the HDSO administrator with their contact details. Patients were then telephoned by the HDSO administrator who collected baseline demographic data (which was pseudonymized and added to the data collected automatically by the HDSO program) and created a username and password for the program. The HDSO administrator also confirmed whether they were happy to take part in research interviews and securely sent SP the ID numbers of all the patients who agreed. The username and password for the HDSO program were then emailed to the patient, along with information on who to contact if there were any problems.

We used qualitative telephone interviews with patients to explore their experience of using the program. We took an HCI approach of rapid data collection and used the data to inform optimization before further evaluation.

One of the members of the HDSO team (SP) contacted patients who registered for the HDSO program but did not start or complete it. We were unable to contact patients who did not register for the program, as they did not provide us with their details or consent to be contacted. The patients who did not start or complete the program were contacted, because we were particularly interested in the problems encountered with the program. The telephone calls were semistructured and lasted approximately 10 minutes. Questions included "What would help you to use the program more regularly?"

The interviews were carried out by telephone by SP, and written notes were taken rather than audio-recording and transcribing because the data needed to be collected and analyzed quickly to inform the program optimization. Note-taking is a recognized form of recording [53], and although it has the disadvantage of not capturing every word verbatim, the researcher mitigated this by noting down verbatim quotes where they were particularly pertinent.

Ethical Approval

Ethical approval was obtained from the HRA (reference number: 159488). Data on registrations, completed sessions, and questionnaire scores were excluded from the HRA REC review, because of a clause that states that secondary analysis of information collected as part of normal care is excluded from REC review by the HRA, as long as the patients are not identifiable [52].

Results

Phase 1: Initial Intervention Components and Content

Establishing User Requirements for HeLP-Diabetes as a Precursor for HDSO

Results from the focus groups showed that patients needed help in managing the complexities of living with diabetes, such as managing the impact that irregular working hours had on diet and blood sugar, impact on relationships and social life, and support in dealing with the profound negative emotions caused by the diagnosis, which included anger, guilt, shame, and despair. The tools to help them manage these tasks included high quality, detailed information, personal stories from other people with similar experiences, and quizzes to test knowledge and provide feedback. Health professionals had similar perceptions of patient needs. Both patients and health professionals wanted HeLP-Diabetes to be interactive and visual (with quizzes, videos, and images), to be easy to use, and have a positive tone [30].

These results were combined with the results of the usability testing to create HeLP-Diabetes, a website containing 560 pages of information divided into 8 sections, which patients at any stage of their illness journey, could dip in and out of.

Content, Structure, and Format of HDSO

We used selected content from the HeLP-Diabetes website (informed by the causal modeling process) to construct the HDSO-structured program, which was needed to meet the QISMET and QOF requirements described in the introduction.

The causal modeling process helped us postulate that information and BCTs targeting healthy eating, weight loss, activity levels, smoking, alcohol consumption, and medication intake, would help users achieve the intended outcomes. Modules targeting these behaviors were therefore selected from the HeLP-Diabetes website to be integrated into the HDSO. The modules contained BCTs, including goal-setting, action-planning, self-monitoring, and feedback on performance [31]. These BCTs are based on the self-regulation theory [54], which states that our major self-regulative mechanism functions through (1) the self-monitoring of behavior, its determinants, and its effects; (2) the judgment of behavior concerning the person and place; and (3) effective self-reaction.

In addition to the BCTs from HeLP-Diabetes, self-assessment questionnaires and feedback were added as new components to the HDSO-structured course. The questionnaires assessed self-efficacy (self-confidence) in self-management, diabetes-related distress, and diabetes knowledge. These were positioned in the course in weeks 1 and 8 (before and after the program), thereby allowing users to reflect on the change in their scores.

Personalized emails were also added as new components to the HDSO-structured course. These were added to encourage motivation and engagement. A systematic review by Alkhaldi et al [55] on the effectiveness of prompts to increase digital interventions found that studies reported borderline small-to-moderate positive effects of technological strategies, including emails, to improve the use of interventions. Resource implications and mindfulness of our ultimate goal being HDSO delivered at scale across the NHS meant that emails were chosen as a cost and time-effective strategy for providing users with reminders.

A curriculum was needed to structure the content and components of the program and to achieve accreditation as a structured course. By identifying relevant theory (as suggested by the MRC guidelines), we decided that the program would follow a spiral curriculum based on the Harden and Stamper spiral curriculum model [56]. This model proposes that there should be an "iterative revisiting of topics, subjects or themes

throughout the course." The idea is that topics are not just repeated, but that knowledge and understanding should be deepened each time. The learner's competence should increase with each visit until the overall aim is achieved [56].

The qualitative work that was conducted to establish user requirements for HeLP-Diabetes showed that users wanted information to be presented using text, images, and videos. These formats were therefore used to present information in the HDSO program and included videos of others living with diabetes. The text was written for people with a reading age of 12 to correspond with 80% of the population in the United Kingdom [57].

The result was an 8-session program containing information presented as text, images, and videos, and BCTs including goal-setting, action-planning, self-monitoring, and feedback on performance. Each session was designed to take approximately 40 to 50 minutes to complete and for people to complete 1 session per week. A screenshot of the HDSO program showing a video giving an introduction to T2DM is shown in Figure 4. The 8 sessions of the program and each of their parts are listed in Table 2.

Figure 4. Screenshot of the HeLP-Diabetes: Starting Out program showing video component.

Table 2. HeLP-Diabetes: Starting Out session titles and parts before usability testing and in the wild testing.

Session title and content	Session components
Week 1—getting started	
Self-assessment	Self-assessment questionnaires
An introduction to diabetes	Information
Eating well for diabetes	Information
Week 2—self-management	
Taking control	Information and quizzes
Becoming more active	Information, physical activity goal-setting task, and videos of people's stories
Handling feelings	Information and videos of people with diabetes
Week 3—improving my health and well-being	
Protecting my body and mind	Information
Making changes	An exercise for reflecting on the quizzes in week 2, and setting SMART[a] goals for diet, medication, activity, drinking, and other health behavior changes
Understanding my moods	Videos of people with diabetes
Working with diabetes	Information
Week 4—taking control of my diabetes	
Making the most of the National Health Service	Information, videos of people talking about their interaction with the National Health Service, and a link to the health record in HeLP-Diabetes where users can record appointments
Update my goals and plans	A review of SMART goals set in week 3
Managing my moods	A reflection on the results of the mood quizzes in week 3, and a set of "Mood Tools," including "Living Life to the Full," a package developed by clinical psychologists using principles from cognitive behavioral therapy
My social life	Information and videos of people's stories
Week 5—medication and lifestyle	
Medication	Information, videos about the challenges and benefits of medications, and an interactive "My medicines" list
Review my goals and plans	A review and update of goals set in week 3
How to fix almost everything	An opportunity to revisit the mood tools used in week 4
Driving	Information
Week 6—reducing my risks	
Reducing the risks of heart attack and strokes	Information
Looking after my feet	Information
Review my goals and plans	An opportunity to review and update SMART goals
Living with diabetes	Videos of people talking about how they became used to having diabetes and an opportunity to revisit mood tools
Week 7—working with my health care team	
Managing illness	Information
My diabetes review	Videos about people's experiences of diabetes care
Review my goals and plans	Review and update SMART goals
Week 8—celebrating success and planning for the future	
Self-assessment	Opportunity to repeat the self-assessment questionnaires from week 1
Looking after my diabetes	Opportunity to prepare a care plan
Moving on: the end of the beginning	Information about staying motivated and reading about diabetes in the media

[a]SMART: specific, measurable, achievable, realistic, and time bound.

Phase 2: Usability Testing

Responses to the questionnaires described earlier were reviewed by the team. The relevant changes were made where there was agreement among responses from the patient volunteers. Where there was disagreement in responses from the patient volunteers, the suggested changes were discussed among the team members, and a consensus decision was made as to whether to make the changes. Changes were made to the language and layout of each session of the program, including making text and titles clearer and easier to understand for users.

Examples of questionnaire responses and changes made to the program, as a result, are given in Table 3, demonstrating how the progression of the development of the program was grounded in user needs.

Table 3. Example usability questionnaire responses and resulting design solutions.

Timeline	Questionnaire item	Response from patient volunteer	Design solutions
Week 1	Design	"Subtitles detailing content for each section would be helpful eg. Self-management 1 - Taking Control; Self-Management 2 - Getting Physical (Becoming more active); Self-management 3 - Handling Emotions (managing feelings?)"	Subtitles were added for each section of each module.
Week 2	Content	"First 3 pages good with 2 very useful video clips. Page 4 'Advice about increasing physical activity' too much detail. Too many peoples stories at the end."	Advice about physical activity was condensed, and the number of videos of people's stories was reduced.
Week 3	Content	"Good, perhaps too much detail (contacts, addresses etc.) for sexual problems - could this be a website link? Level of detail might be off-putting for newly diagnosed."	Contact details for support and advice organizations were removed and website links were added and signposted instead.

We made changes to the design and content of the program based on the questionnaire responses. This ensured that the development of the program was grounded in user needs. The changes included clearer subtitling, advice about physical activity made more concise, a more appropriate number of videos of people's stories included, and links to support and advice organizations added.

Phase 3: In the Wild Testing

During the study, 24 people registered for the HDSO program. Quantitative data were collected on program use, questionnaire scores, and characteristics. Of the 24 people registered, 3 (13%) people completed the program, 13 (54%) people started the program but did not complete it, and 8 (33%) people did not start the program. The data suggested low uptake and completion, but those who used it seemed to benefit from it (mean self-efficacy in self-management scores and diabetes knowledge scores increased).

The telephone interview responses were analyzed using a thematic analysis approach, and a list of barriers to completing the program emerged from the data. These are listed in Table 4 with illustrative quotes.

Table 4. Themes observed from the telephone interviews.

Theme	Illustrative quote
Lack of time to start or complete the program	• "I've tried going through it during breaks at work, but I keep getting interrupted. I've only got to the 'Welcome' page." • "Can you give me an extra hour in the day?"
Expectation that completing the program would take too long	• "It's going to take a while, I need to be able to use it with a spare ten minutes."
Ambivalence about starting	• "It's in the background, I keep it in mind."
Feeling of content not being relevant to some users	• "It's not relevant to me, I don't take medication."

These themes were used to inform the refinement and optimization of the program as discussed in the next section. The ideas were followed up in subsequent interviews conducted as part of the evaluation of the final intervention and reported elsewhere [34].

Design Solutions Resulting From the In the Wild Testing

The themes identified from the interviews suggested that there were patient and program factors which influenced program use. Patient factors such as ambivalence were difficult to address. However, we were able to shorten the program and provide users with quicker access to the program with web-based registration.

Following a discussion among the HDSO team, the following changes were agreed upon:

1. Reducing the number of sessions in the program: evidence from systematic reviews of engagement with digital behavior change interventions [58] and research on adult web-based learning [59] suggests that participants disengage if the intervention is perceived as too long or overly complicated. The decision about what content to retain and what to remove was made after discussions with the diabetes

specialist nurses in the HDSO team who were trained educators and experienced in delivering face-to-face structured education courses. To determine what content to retain and what to remove, we discussed the data from the user experience interviews, reviewed the guidelines on T2DM management [26], and examined the curriculum closely. The 8 sessions were cut down to 4 sessions, with a fifth bonus session available at the end. Despite comments about its irrelevance to newly diagnosed patients, all aspects of the management of T2DM were retained because it was considered important to give people a good overview and understanding of the types of treatment they might receive in the future. Topics including *managing my diabetes when I'm ill*, *working with diabetes*, and *driving with diabetes* were taken out of the main course and moved to the fifth bonus session. The final 4-session intervention is described elsewhere [60] and contains the following sessions: *getting started*, *self-management*, *improving my health and well-being*, and *taking control of my diabetes*.

2. Reducing the number of questionnaires: we decided to reduce the number of questionnaires from 3 to 2 by removing the AdKnowl questionnaire. The AdKnowl (knowledge) questionnaire [61] was removed because it was significantly longer and more time-consuming than the other 2 questionnaires. Patient feedback suggested that they found the questionnaire burdensome and off-putting; the evidence we found from systematic reviews of diabetes self-management education programs suggests that there is a lack of a consistent positive relationship between knowledge and glycemic control and that factors other than knowledge are needed to achieve long-term behavior change [21]. Therefore, we prioritized the changes in distress and self-efficacy.
3. Web-based self-registration: it was decided to change to web-based registration to save time and to make it easier for patients to access the program quickly. The self-registration page included a demographic questionnaire, which allowed for the collection of baseline data. Telephone support from the HDSO team was still available for those who had difficulty registering on the web or using the program.

We decided to offer the program to everyone with T2DM and not just people who were newly diagnosed. The HDSO program was developed in line with the national clinical guidelines for GPs advising them to offer patients with T2DM structured education at and around the time of diagnosis [26]. However, we knew from the National Diabetes Audit that not all patients were offered structured education at the time of diagnosis, and of those who were offered it in 2016-2017, only 7.1% attended [29]. Therefore, many patients with T2DM who were not newly diagnosed have not received structured self-management education and are in need of it. In addition, data on the incidence and prevalence of T2DM in the United Kingdom show that T2DM prevalence rates have more than doubled between 2000 and 2013, but incidence rates have increased more slowly [62,63]. This suggests that there are more people being diagnosed younger and living longer rather than new diagnoses, which consequently suggests that it would be possible to recruit people who were not newly diagnosed to the HDSO program than people who were newly diagnosed. We decided to offer the program to everyone with T2DM and collect data on the duration since diagnosis. This allowed us to compare completion rates between newly diagnosed and non–newly diagnosed patients.

Discussion

Principal Findings

We have described the stages of the development of a web-based structured education program for people newly diagnosed with T2DM, HDSO. Methods from HCI and health research were used in combination at every stage. Methods from HCI put more emphasis on understanding user requirements and determining uptake. The *in the wild* testing allowed us to identify low completion rates, which were not picked up in the usability testing because it was conducted with highly motivated patient volunteers. The methods from health and the MRC framework for complex interventions emphasized the impact of the intervention. This helped us to understand the potential effectiveness of the intervention. The iterative development process that we went through with the intervention, in contrast to traditional piloting and feasibility studies conducted in health research [64], resulted in an intervention that was more stable and appropriate for a definitive trial than the earlier iteration. Proceeding to a trial too early can be problematic because trials do not detect whether the lack of intervention effect is due to implementation failure or genuine ineffectiveness [65]. Randomized controlled trials also fail to permit iterative improvements to the design and updates to technology [66].

Comparison With Previous Work

The existing reporting of complex behavior change interventions is limited, and this prevents successful replication of successful interventions [14]. Reviews of web-based T2DM self-management interventions have reported extreme heterogeneity of interventions [67] and poor descriptions of the theoretical bases and active ingredients of the interventions [39]. This makes it difficult for researchers to identify and understand successful intervention components and to be able to design and implement successful interventions. The field of digital health research and web-based diabetes self-management is evolving rapidly, and it is important for future research that lessons can be learned from existing studies. This description of the development and content of HDSO helps add to the understanding of how and why web-based interventions for diabetes self-management (and other long-term conditions) work and can be used to inform future research in this area. In addition to describing the intervention, this study also adds to the understanding of how interdisciplinary methods from health and HCI can be used to develop a DHI. Previous studies by Blandford et al [2] and Pagliari [11] have described the challenges in using interdisciplinary research in the development and evaluation of DHIs, and this paper illustrates some of the concepts described in the literature using the example of a web-based diabetes self-management program.

Strengths and Limitations

The strength of this research is the use of methods from both health and HCI. We combined the theory of living with long-term conditions and a user-centered design to understand and meet user requirements. This meant that users of different age groups, education levels, and ethnic backgrounds could use the intervention, as demonstrated by the analysis of the usage data from subsequent studies [68]. A weakness of the interdisciplinary approach in this study was the emphasis on time and resources when conducting qualitative interviews with program users. This meant that we conducted rapid user experience studies at several stages, with smaller sample sizes and limited depth and duration of interviews.

Another weakness of our approach was patient involvement. More extensive patient involvement could have been used in the design of the interview guide for the telephone interviews in the user experience study and data analysis. The patient volunteers could also have been asked to make their own suggestions for refinements to the first iteration of the program, instead of relying solely on the data.

Conclusions

This paper describes how interdisciplinary methods can be used to develop a web-based structured education program for people newly diagnosed with T2DM. Methods were combined from human-computer research and health research. The reporting of the development processes for DHIs needs to continue, especially when interdisciplinary methods are used, for researchers to be able to learn from each other and create user-centered interventions.

Acknowledgments

The authors would like to thank the participants and professionals who contributed to the development of the HeLP-Diabetes: Starting Out program, particularly Orla O'Donnell, Helen Gibson, and Rebecca Owen.

Funding for a percentage of the website development and a percentage of hosting and maintenance of the website was provided by the National Institute for Health Research School for Primary Care Research (project reference: 280).

The views expressed are those of the authors and not necessarily those at the National Institute for Health Research or the Department of Health and Social Care.

Authors' Contributions

SP, FH, and EM contributed to this research proposal. EM, SP, JR, and KP were involved in the development of the intervention. SP obtained ethical approval, recruited patients, collected and analyzed the data, and drafted the manuscript. All authors, including NN, reviewed and edited the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

EM was the Managing Director of the HeLP-Digital Community Interest Company during the period of the research, a not-for-profit social enterprise established to roll out HeLP-Diabetes. She was not remunerated for this role, and there was no financial gain from the publication of this paper.

References

1. 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. Br Med J 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
2. Blandford A, Gibbs J, Newhouse N, Perski O, Singh A, Murray E. Seven lessons for interdisciplinary research on interactive digital health interventions. Digit Health 2018 May 03;4:2055207618770325 [FREE Full text] [doi: 10.1177/2055207618770325] [Medline: 29942629]
3. Elwyn G, Kreuwel I, Durand MA, Sivell S, Joseph-Williams N, Evans R, et al. How to develop web-based decision support interventions for patients: a process map. Patient Educ Couns 2011 Feb;82(2):260-265. [doi: 10.1016/j.pec.2010.04.034] [Medline: 20627644]
4. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2011; Vancouver BC Canada p. 3063-3072. [doi: 10.1145/1978942.1979396]
5. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
6. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: new guidance 2006. Medical Research Council. 2006. URL: https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/ [accessed 2021-12-27]
7. Dix A, Finlay J, Abowd G, Beale R. Human-Computer Interaction. New York: Pearson; 2004.

8. Rogers Y, Sharp H, Preece J. Interaction Design: Beyond Human-Computer Interaction Third Edition Ed. Hoboken, New Jersey, United States: John Wiley & Sons; 2013.
9. Calp MH, Akcayol MA. The importance of human-computer interaction in the development process of software projects. Int J Inform Technol Emerg Technol 2015 Nov 25:48. [doi: 10.18844/gjit.v0i0.114]
10. Maguire M. Methods to support human-centred design. Int J Hum-Comput Stud 2001 Oct;55(4):587-634. [doi: 10.1006/ijhc.2001.0503]
11. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007 May 27;9(2):e15 [FREE Full text] [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]
12. Editorial. Current issues and future directions in psychology and health: towards a cumulative science of behaviour change: Do current conduct and reporting of behavioural interventions fall short of best practice? Psychol Health 2007 Dec;22(8):869-874. [doi: 10.1080/08870440701520973]
13. Riley BL, MacDonald J, Mansi O, Kothari A, Kurtz D, vonTettenborn LI, et al. Is reporting on interventions a weak link in understanding how and why they work? A preliminary exploration using community heart health exemplars. Implement Sci 2008 May 20;3(1):27 [FREE Full text] [doi: 10.1186/1748-5908-3-27] [Medline: 18492247]
14. Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implement Sci 2009 Jul 16;4(1):40 [FREE Full text] [doi: 10.1186/1748-5908-4-40] [Medline: 19607700]
15. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
16. O'Cathain A, Croot L, Sworn K, Duncan E, Rousseau N, Turner K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. Pilot Feasibility Stud 2019 Mar 12;5(1):41 [FREE Full text] [doi: 10.1186/s40814-019-0425-6] [Medline: 30923626]
17. O'Brien N, Heaven B, Teal G, Evans EH, Cleland C, Moffatt S, et al. Integrating evidence from systematic reviews, qualitative research, and expert knowledge using co-design techniques to develop a web-based intervention for people in the retirement transition. J Med Internet Res 2016 Aug 03;18(8):e210 [FREE Full text] [doi: 10.2196/jmir.5790] [Medline: 27489143]
18. White BK, Martin A, White JA, Burns SK, Maycock BR, Giglia RC, et al. Theory-based design and development of a socially connected, gamified mobile app for men about breastfeeding (Milk Man). JMIR Mhealth Uhealth 2016 Jun 27;4(2):e81 [FREE Full text] [doi: 10.2196/mhealth.5652] [Medline: 27349756]
19. Curtis KE, Lahiri S, Brown KE. Targeting parents for childhood weight management: development of a theory-driven and user-centered healthy eating app. JMIR Mhealth Uhealth 2015 Jun 18;3(2):e69 [FREE Full text] [doi: 10.2196/mhealth.3857] [Medline: 26088692]
20. Deakin T, McShane CE, Cade JE, Williams RD. Group based training for self-management strategies in people with type 2 diabetes mellitus. Cochrane Database Syst Rev 2005 Apr 18(2):CD003417. [doi: 10.1002/14651858.CD003417.pub2] [Medline: 15846663]
21. Norris SL, Engelgau MM, Narayan KM. Effectiveness of self-management training in type 2 diabetes: a systematic review of randomized controlled trials. Diabetes Care 2001 Mar 01;24(3):561-587. [doi: 10.2337/diacare.24.3.561] [Medline: 11289485]
22. Ellis SE, Speroff T, Dittus RS, Brown A, Pichert JW, Elasy TA. Diabetes patient education: a meta-analysis and meta-regression. Patient Edu Counsel 2004 Jan;52(1):97-105. [doi: 10.1016/s0738-3991(03)00016-8]
23. Steinsbekk A, Rygg L, Lisulo M, Rise MB, Fretheim A. Group based diabetes self-management education compared to routine treatment for people with type 2 diabetes mellitus. A systematic review with meta-analysis. BMC Health Serv Res 2012 Jul 23;12(1):213 [FREE Full text] [doi: 10.1186/1472-6963-12-213] [Medline: 22824531]
24. Diabetes Prevalence Model 2016. Public Health England. URL: https://www.diabetes.org.uk/professionals/position-statements-reports/statistics/diabetes-prevalence-2016 [accessed 2021-12-27]
25. Hex N, Bartlett C, Wright D, Taylor M, Varley D. Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. Diabet Med 2012 Jul;29(7):855-862. [doi: 10.1111/j.1464-5491.2012.03698.x] [Medline: 22537247]
26. Type 2 Diabetes in adults: management. NICE guideline NG28. National Institute for Health and Care Excellence. 2015. URL: https://www.nice.org.uk/guidance/ng28 [accessed 2021-01-31]
27. DSME Standard. Quality Institute for Self-Management Education. 2010. URL: https://www.qismet.org.uk/ [accessed 2021-02-14]
28. 2016-17 General Medical Services (GMS) contract and Quality and Outcomes Framework (QOF). NHS Digital. 2017. URL: https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/quality-and-outcomes-framework-qof-2016-17 [accessed 2021-12-27]
29. National Diabetes Audit, 2017-18 : Care Processes and Treatment Targets short report. NHS Digital. 2018. URL: https://files.digital.nhs.uk/E6/369B83/

National%20Diabetes%20Audit%202017-18%20Short%20Report%2C%20Care%20Processes%20and%20Treatment%20Targets.pdf [accessed 2021-12-27]
30. Dack C, Ross J, Stevenson F, Pal K, Gubert E, Michie S, et al. A digital self-management intervention for adults with type 2 diabetes: combining theory, data and participatory design to develop HeLP-Diabetes. Internet Interv 2019 Sep;17:100241 [FREE Full text] [doi: 10.1016/j.invent.2019.100241] [Medline: 31372349]
31. Murray E. Development, evaluation and implementation of a computer-based self-management programme for people with type 2 diabetes. NIHR Journals Library. 2017. URL: https://www.journalslibrary.nihr.ac.uk/programmes/pgfar/RP-PG-0609-10135/#/ [accessed 2021-12-30]
32. Li J, Parrott S, Sweeting M, Farmer A, Ross J, Dack C, et al. Cost-effectiveness of facilitated access to a self-management website, compared to usual care, for patients with type 2 diabetes (HeLP-Diabetes): randomized controlled trial. J Med Internet Res 2018 Jun 08;20(6):e201 [FREE Full text] [doi: 10.2196/jmir.9256] [Medline: 29884608]
33. Alkhaldi G, Modrow K, Hamilton F, Pal K, Ross J, Murray E. Promoting engagement with a digital health intervention (HeLP-Diabetes) using email and text message prompts: mixed-methods study. Interact J Med Res 2017 Aug 22;6(2):e14 [FREE Full text] [doi: 10.2196/ijmr.6952] [Medline: 28829328]
34. Poduval S, Marston L, Hamilton F, Stevenson F, Murray E. Feasibility, acceptability, and impact of a web-based structured education program for type 2 diabetes: real-world study. JMIR Diabetes 2020 Jan 06;5(1):e15744 [FREE Full text] [doi: 10.2196/15744] [Medline: 31904580]
35. Lazar J. Web Usability: A User-Centered Design Approach. Boston, MA: Addison-Wesley; 2006.
36. Rubin J, Chisnell D, Spool J. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. Hoboken, New Jersey, United States: John Wiley & Sons; 2008.
37. Yen P, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
38. Nielsen J, Clemmensen T, Yssing C. Getting access to what goes on in people's heads?: reflections on the think-aloud technique. In: Proceedings of the Second Nordic Conference on Human-computer Interaction. 2002 Presented at: NORDICHI02: NORDICHI 2002, Proceedings of the Second Nordic Conference on Human-computer Interaction; October 19 - 23, 2002; Aarhus Denmark p. 101-110. [doi: 10.1145/572020.572033]
39. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013 Mar 28(3):CD008776 [FREE Full text] [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
40. Pereira K, Phillips B, Johnson C, Vorderstrasse A. Internet delivered diabetes self-management education: a review. Diabetes Technol Ther 2015 Jan;17(1):55-63. [doi: 10.1089/dia.2014.0155] [Medline: 25238257]
41. Ramadas A, Quek K, Chan C, Oldenburg B. Web-based interventions for the management of type 2 diabetes mellitus: a systematic review of recent evidence. Int J Med Inform 2011 Jun;80(6):389-405. [doi: 10.1016/j.ijmedinf.2011.02.002] [Medline: 21481632]
42. van Vugt M, de Wit M, Cleijne WH, Snoek FJ. Use of behavioral change techniques in web-based self-management programs for type 2 diabetes patients: systematic review. J Med Internet Res 2013 Dec 13;15(12):e279 [FREE Full text] [doi: 10.2196/jmir.2800] [Medline: 24334230]
43. Corbin JM, Strauss A. Unending Work and Care: Managing Chronic Illness at Home. San Francisco: Jossey-Bass; 1988.
44. Merriam SB. Adult learning theory for the twenty-first century. New Directions for Adult and Continuing Education 2008 Jun;2008(119):93-98. [doi: 10.1002/ace.309]
45. Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/s15324796abm2601_01]
46. Michie S, Carey RN, Johnston M, Rothman AJ, de Bruin M, Kelly MP, et al. From theory-inspired to theory-based interventions: a protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. Ann Behav Med 2018 May 18;52(6):501-512 [FREE Full text] [doi: 10.1007/s12160-016-9816-6] [Medline: 27401001]
47. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
48. Crabtree A, Chamberlain A, Grinter RE, Jones M, Rodden T, Rogers Y. Introduction to the special issue of “The Turn to The Wild”. ACM Trans Comput-Hum Interact 2013 Jul;20(3):1-4. [doi: 10.1145/2491500.2491501]
49. Rogers Y, Marshall P. Research in the wild. Synth Lect Hum-Cent Inform 2017 Apr 04;10(3):i-97. [doi: 10.2200/s00764ed1v01y201703hci037]
50. Snoek F, Welch G. Problem Areas in Diabetes (PAID) questionnaire in DAWN (Diabetes Attitudes, Wishes and Needs) study. Novo Nordisk. 2006. URL: https://www.huskyhealthct.org/providers/provider_postings/diabetes/PAID_problem_areas_in_diabetes_questionnaire.pdf [accessed 2021-12-27]
51. Sturt J, Hearnshaw H, Wakelin M. Validity and reliability of the DMSES UK: a measure of self-efficacy for type 2 diabetes self-management. Primary Health Care 2010 Jun 22;11(04):374-381. [doi: 10.1017/s1463423610000101]

52. What approvals and decisions do I need? NHS Health Research Authority. 2016. URL: http://www.hra.nhs.uk/resources/before-you-apply/research-requiring-nhs-rd-review-but-not-ethical-review/ [accessed 2021-12-27]
53. Pope C, Mays N. Qualitative Research in Health Care, Third Edition. Hoboken, New Jersey, United States: Blackwell Publishing Ltd; 2006.
54. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Dec Proc 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-l]
55. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 08;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]
56. Harden RM. What is a spiral curriculum? Med Teach 1999 Jul 03;21(2):141-143. [doi: 10.1080/01421599979752] [Medline: 21275727]
57. Adult Literacy 2017. National Literacy Trust. URL: https://literacytrust.org.uk/parents-and-families/adult-literacy/ [accessed 2021-12-27]
58. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
59. Jordan K. Massive open online course completion rates revisited: assessment, length and attrition. Int Rev of Res Open Distrib Learn 2015 Jun 19;16(3). [doi: 10.19173/irrodl.v16i3.2112]
60. Poduval S, Ross J, Pal K, Newhouse N, Hamilton F, Murray E. Use of the TIDieR checklist to describe an online structured education programme for type 2 diabetes. Digit Health 2020 Nov 30;6:2055207620975647 [FREE Full text] [doi: 10.1177/2055207620975647] [Medline: 33294208]
61. Speight JC, Bradley C. The ADKnowl: identifying knowledge deficits in diabetes care. Diabet Med 2001 Aug;18(8):626-633. [doi: 10.1046/j.1464-5491.2001.00537.x] [Medline: 11553199]
62. Zghebi SS, Steinke DT, Carr MJ, Rutter MK, Emsley RA, Ashcroft DM. Examining trends in type 2 diabetes incidence, prevalence and mortality in the UK between 2004 and 2014. Diabetes Obes Metab 2017 Nov 05;19(11):1537-1545. [doi: 10.1111/dom.12964] [Medline: 28387052]
63. Sharma M, Nazareth I, Petersen I. Trends in incidence, prevalence and prescribing in type 2 diabetes mellitus between 2000 and 2013 in primary care: a retrospective cohort study. BMJ Open 2016 Jan 13;6(1):e010210 [FREE Full text] [doi: 10.1136/bmjopen-2015-010210] [Medline: 26769791]
64. Feasibility and pilot studies. National Institute for Health Research. 2015. URL: https://www.nihr.ac.uk/funding-and-support/documents/funding-for-research-studies/research-programmes/PGfAR/CCF-PGfAR-Feasibility-and-Pilot-studies.pdf [accessed 2021-12-27]
65. Levati S, Campbell P, Frost R, Dougall N, Wells M, Donaldson C, et al. Optimisation of complex health interventions prior to a randomised controlled trial: a scoping review of strategies used. Pilot Feasibility Stud 2016 Mar 15;2(1):17 [FREE Full text] [doi: 10.1186/s40814-016-0058-y] [Medline: 27965837]
66. Mohr DC, Schueller SM, Riley WT, Brown CH, Cuijpers P, Duan N, et al. Trials of intervention principles: evaluation methods for evolving behavioral intervention technologies. J Med Internet Res 2015 Jul 08;17(7):e166 [FREE Full text] [doi: 10.2196/jmir.4391] [Medline: 26155878]
67. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. J Diabetes Sci Technol 2017 Sep 31;11(5):1015-1027 [FREE Full text] [doi: 10.1177/1932296817713506] [Medline: 28560898]
68. Poduval S, Ahmed S, Marston L, Hamilton F, Murray E. Crossing the digital divide in online self-management support: analysis of usage data from help-diabetes. JMIR Diabetes 2018 Dec 06;3(4):e10925 [FREE Full text] [doi: 10.2196/10925] [Medline: 30522988]

Abbreviations

BCT: behavior change technique
DHI: digital health intervention
GP: general practitioner
HCI: human-computer interaction
HDSO: HeLP-Diabetes: Starting Out
HRA: Health Research Authority
MRC: Medical Research Council
NHS: National Health Service
QISMET: Quality Institute for Self-Management Education
QOF: Quality and Outcomes Framework
REC: Research Ethics Committee
T2DM: type 2 diabetes mellitus

Edited by A Kushniruk; submitted 25.06.21; peer-reviewed by D Lungu, E Børøsund; comments to author 26.08.21; revised version received 24.09.21; accepted 15.10.21; published 14.01.22.

Please cite as:
Poduval S, Ross J, Pal K, Newhouse N, Hamilton F, Murray E
Web-Based Structured Education for Type 2 Diabetes: Interdisciplinary User-Centered Design Approach
JMIR Hum Factors 2022;9(1):e31567
URL: https://humanfactors.jmir.org/2022/1/e31567
doi:10.2196/31567
PMID:35029531

Original Paper

Developing a User-Centered Digital Clinical Decision Support App for Evidence-Based Medication Recommendations for Type 2 Diabetes Mellitus: Prototype User Testing and Validation Study

Kevin Larsen[1], MD; Bilikis Akindele[2], MD, MMCi; Henry Head[3], BSc; Rick Evans[3], BSc; Purvi Mehta[3], MD, MBI; Quinn Hlatky[3], BBA; Brendan Krause[4], BA, MSc; Sydney Chen[3], BSc; Dominic King[3], MD, PhD

[1]Center for Advanced Clinical Solution, OptumHealth, Optum, Washington, DC, United States
[2]Center for Advanced Clinical Solution, OptumHealth, Optum, Boston, MA, United States
[3]Center for Advanced Clinical Solution, OptumHealth, Optum, Raleigh, NC, United States
[4]OptumLabs, UnitedHealth Group, Minnetoka, MN, United States

Corresponding Author:
Bilikis Akindele, MD, MMCi
Center for Advanced Clinical Solution
OptumHealth
Optum
1325 Boylston St
Boston, MA, 02215
United States
Phone: 1 984 227 2621
Email: bilikis.akindele@optum.com

Abstract

Background: Closing the gap between care recommended by evidence-based guidelines and care delivered in practice is an ongoing challenge across systems and delivery models. Clinical decision support systems (CDSSs) are widely deployed to augment clinicians in their complex decision-making processes. Despite published success stories, the poor usability of many CDSSs has contributed to fragmented workflows and alert fatigue.

Objective: This study aimed to validate the application of a user-centered design (UCD) process in the development of a standards-based medication recommender for type 2 diabetes mellitus in a simulated setting. The prototype app was evaluated for effectiveness, efficiency, and user satisfaction.

Methods: We conducted interviews with 8 clinical leaders with 8 rounds of iterative user testing with 2-8 prescribers in each round to inform app development. With the resulting prototype app, we conducted a validation study with 43 participants. The participants were assigned to one of two groups and completed a 2-hour remote user testing session. Both groups reviewed mock patient facts and ordered diabetes medications for the patients. The Traditional group used a mock electronic health record (EHR) for the review in Period 1 and used the prototype app in Period 2, while the Tool group used the prototype app during both time periods. The perceived cognitive load associated with task performance during each period was assessed with the National Aeronautics and Space Administration Task Load Index. Participants also completed the System Usability Scale (SUS) questionnaire and Kano Survey.

Results: Average SUS scores from the questionnaire, taken at the end of 5 of the 8 user testing sessions, ranged from 68-86. The results of the validation study are as follows: percent adherence to evidence-based guidelines was greater with the use of the prototype app than with the EHR across time periods with the Traditional group (prototype app mean 96.2 vs EHR mean 72.0, $P<.001$) and between groups during Period 1 (Tool group mean 92.6 vs Traditional group mean 72.0, $P<.001$). Task completion times did not differ between groups ($P=.23$), but the Tool group completed medication ordering more quickly in Period 2 (Period 1 mean 130.7 seconds vs Period 2 mean 107.7 seconds, $P<.001$). Based on an adjusted α level owing to violation of the assumption of homogeneity of variance (Ps>.03), there was no effect on screens viewed and on perceived cognitive load (all Ps>.14).

Conclusions: Through deployment of the UCD process, a point-of-care medication recommender app holds promise of improving adherence to evidence-based guidelines; in this case, those from the American Diabetes Association. Task-time performance suggests that with practice the T2DM app may support a more efficient ordering process for providers, and SUS scores indicate provider satisfaction with the app.

(JMIR Hum Factors 2022;9(1):e33470) doi:10.2196/33470

KEYWORDS

clinical decision support; user-centered design; user testing; type 2 diabetes mellitus; evidence-based guidelines; validation; workflows; electronic health record; decision support; design; diabetes

Introduction

Background

Diabetes affects roughly 34.2 million Americans, 90-95% of whom have type 2 diabetes mellitus (T2DM) [1]. Another 88 million adults in the United States have a condition called prediabetes, which puts them at risk for T2DM [1]. In addition to the quality-of-life challenges associated with managing the disease, T2DM can be associated with an array of complications, including kidney failure, blindness, and amputation of a toe, foot, or leg [2]. Every year, an estimated US $237 billion of the health care budget is spent on treating and managing the disease [2].

The high costs and suboptimal outcomes associated with T2DM may be associated at least in part with variability of care. For example, studies have shown that maintaining goal glycated hemoglobin (HbA_{1c}) values can prevent or delay diabetes-related complications and decrease direct medical costs [2,3]. However, studies demonstrate significant variability in care paths for people diagnosed with T2DM [4], despite existing guidelines for specific lines of care; this negatively impacts their ability to achieve key health outcomes. This gap between care recommended by evidence-based guidelines and care delivered in practice is due in part to the sheer volume of information that providers must routinely digest as evidence and recommendations continually evolve.

Over the past decade, point-of-care clinical decision support systems (CDSSs) have emerged as one approach to close this gap. These systems can manifest as order sets, computerized alerts and reminders, digital guidelines, and clinical workflow tools designed to augment and support provider capabilities. Core functions of these software applications include summarizing patient facts, visualizing trends, supporting documentation, generating reminders, and making therapy recommendations to the provider [5,6].

CDSSs generally support the provider by (1) pulling together relevant patient facts in a manner that is efficiently assessed, and (2) bringing up-to-date, evidence-based clinical guidelines to the point of care where clinical decisions are made. Sutton et al [5] identified benefits of adoption along with risks that should be mitigated through strategic design. For example, a CDSS may elevate adherence to clinical guidelines, but it may risk creating excessive trust in the system without appropriate checks. Improved retrieval and presentation of patient data through a CDSS may support better choices in treatment, but it also risks disrupting existing workflow if usability is not adequately evaluated. Determining the success of clinical decision support (CDS) tools ultimately depends on measurable improvements in the quality of care. The literature provides examples of improvements in process-related and clinical outcomes [4]. For example, in randomized trials, CDS interventions have been associated with increased hemoglobin testing rates and with steeper declines in measured HbA_{1c} levels—an indication of glycemic control [6-9].

Measurable quality-of-care improvements are dependent on good CDSS design. Sim et al [10] developed and tested a web-based CDS tool, a diabetes dashboard that provided graphic summarization of laboratory results and was intended to facilitate the interpretation of results and flag tests needed by the patients. User testing demonstrated performance advantages over the electronic health record (EHR) for recognition of abnormal test results, identification of long-term trends, and awareness of which tests were due for repeating. However, participants using the dashboard were not able to better determine whether a treatment adjustment was required. The failure to find a treatment-decision benefit should not be unexpected given that the dashboard was not designed to make patient-specific treatment recommendations. The outcome highlights the need for an interface that supports the processing of patient facts and leverages evidence-based guidelines.

Objectives

The aim of this study was to evaluate the application of a user-centered design (UCD) process toward the development of a prototype software that serves as a medication recommender for T2DM (the T2DM app). The prototype T2DM app is a standards-based CDS tool that provides evidence-based medication recommendations to health care providers with the aim of improving adherence to the latest evidence-based clinical guidance and reducing cognitive load for clinicians making prescribing decisions. The prototype T2DM app integrates into existing EHR systems that clinicians use as part of standard practice to review patient records, order tests and prescriptions, etc.

Methods

Methods Overview

The prototype app development consisted of two phases: (1) a predevelopment analytic phase to learn about user needs, the context of use, and specific workflow associated with reviewing patient facts and ordering medications; and (2) iterative user testing of the prototype app itself. The validation study with the final version of the prototype app addressed the following research questions: (1) is medication ordering with the T2DM app associated with more medication orders that align with American Diabetes Association (ADA) evidence–based guidelines compared to ordering medications with a typical EHR? (2) Is medication ordering with the T2DM app associated with faster overall task times compared to ordering medications with a typical EHR? (3) Is medication ordering with the T2DM app associated with lower perceived cognitive load (as measured by the National Aeronautics and Space Administration Task

Load Index [NASA TLX]) compared to ordering medications with a typical EHR?

Prototype App Development

Analytic Phase

The analytic work that provided the foundation for the first interactive prototype entailed the review of ADA guidelines [11] to determine priorities for the selection of patient facts that would need to be pulled into the T2DM app from the EHR. The team also conducted interviews with 8 conveniently selected clinical leaders from care delivery organizations of significant scale to learn about the intended users, the context of use, and the specific workflow in their EHR associated with reviewing patient facts and ordering medications to manage T2DM.

User Testing Phase

Once the initial prototype was developed on the basis of findings from the analytic phase, 8 iterative rounds of user testing were carried out to get feedback on different parts of its evolving design. These were conducted remotely from May 4 to October 2, 2020, owing to COVID-19 restrictions, and each session was recorded. For each round of user feedback, a sample of prescribers were included in 1-hour, one-on-one test sessions with the prototype app. In total, 16 participants were recruited directly from a single large provider network, and the remaining participants were recruited via a national third-party recruitment service. The latter were compensated at the prevailing market rate for physicians, physician assistants, and nurse practitioners. In total, 25 MD physicians, 11 nurse practitioners, and 15 physician assistants participated in the user testing. We used a talk-aloud method for data collection while prescribers reviewed and ordered medications with the prototype.

In addition, during 5 of the 8 rounds of user testing, participants completed the System Usability Scale (SUS) questionnaire at the end of their sessions (for logistical reasons [sample size and participant time], the SUS was not administered in rounds 5, 7, and 8). The SUS questionnaire is an industry standard for evaluating the usability of software applications consisting of 10 statements with 5 response options (ranging from “Strongly Disagree” to “Strongly Agree”) to each question. The statements are as follows:

- I think that I would like to use this system frequently.
- I found the system unnecessarily complex.
- I thought the system was easy to use.
- I think that I would need the support of a technical person to be able to use this system.
- I found the various functions in the system were well integrated.
- I thought there was too much inconsistency in this system.
- I would imagine that most people would learn to use this system very quickly.
- I found the system very cumbersome to use.
- I felt very confident using the system.
- I needed to learn a lot of things before I could get going with this system.

The SUS survey yields a single number that represents a composite measure of the overall perceived usability of the system. SUS scores have a range of 0 to 100 and the score is a relative benchmark that is used against other iterations of the system. The SUS is a reliable and valid measure of system satisfaction. Sauro [12] reports that the average SUS score from 500 studies across various products (eg, websites, cellphones, and enterprise systems) and across different industries is 68. A SUS score above 68 is considered above average and anything below 68 is below average.

Validation Study

Participants

In total, 43 participants completed the 2-hour remote evaluation study of the prototype T2DM app that was developed and refined through the user testing phase. Participants were recruited via a national third-party recruitment service. The study population included 21 MD and 22 non-MD physicians (5 nurse practitioners, 5 physician assistants, 6 nurses, and 6 pharmacists). To be included in the study, candidates were required to (1) have at least 1 year of experience treating T2DM, preferably in a family or internal medicine practice; (2) currently prescribe, prescribe on behalf of, or provide medication recommendations as part of their current role; and (3) currently interact with 15 or more patients per day. Participant compensation was set at the prevailing market rate for physician assistants, nurse practitioners, pharmacists, and nurses.

Medical and legal review were conducted to ensure no aspect of clinical or legal regulations or ethical considerations were overlooked. Formal institutional review board or ethical review was not required in the study because no protected health information was included, and participation was limited to usability testing and providing feedback about the app. No private information about participants was collected or included in the analysis and study. All participants gave verbal consent and were reimbursed for their time.

A Universal Design framework was followed for accessibility and to accommodate a wide range of people (eg, people with color blindness), and care was taken to convert typical in-person interaction to virtual to avoid exposure to COVID-19.

Stimuli

The prototype T2DM app (Figure 1) pulls data from the EHR to present patient information in a user-friendly way on the left side of the screen. The right side of the screen displays a list of current diabetes medications, evidence-based recommendations based on the latest ADA guidelines, and a table of on- and off-guideline medications for ordering. In addition to the prototype T2DM app, a second prototype that is a close representation of a commercial EHR user interface was developed for use in the validation study (Figure 2). Both prototypes (the mock EHR and the T2DM) presented mock patient facts that simulate patients with T2DM. The workflow was captured as each participant interacted with both prototypes (Figures 3 and 4).

Figure 1. A screenshot of the prototype T2DM app. A1C: glycated hemoglobin, ASCVD: atherosclerotic cardiovascular disease, CHF: congestive heart failure, CKD: chronic kidney disease, eGFR: estimated glomerular filtration rate.

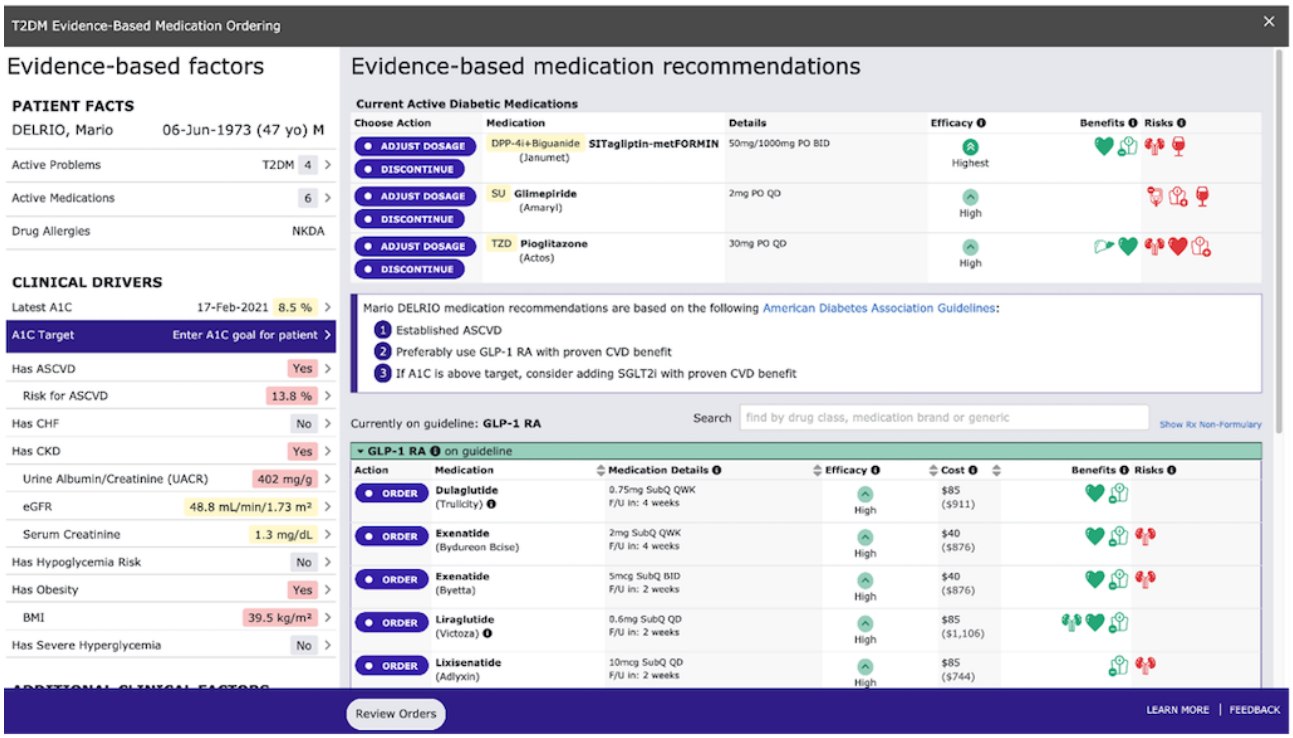


Figure 2. A screenshot of the mock electronic health record. BP: blood pressure.

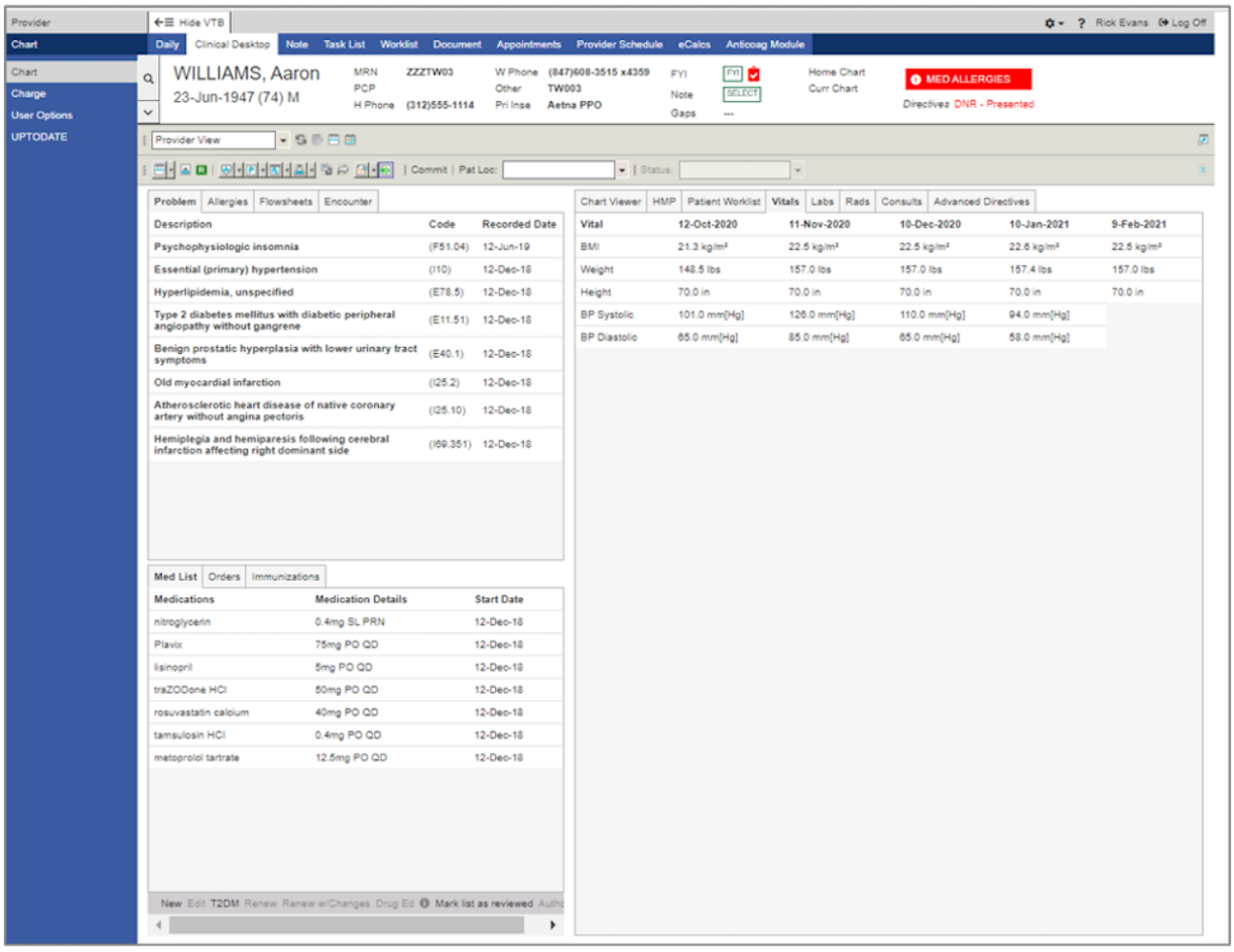


Figure 3. Schematic of workflow to complete an ordering task in an EHR that includes the integrated prototype T2DM app. EHR: electronic health record, T2DM: type 2 diabetes mellitus.

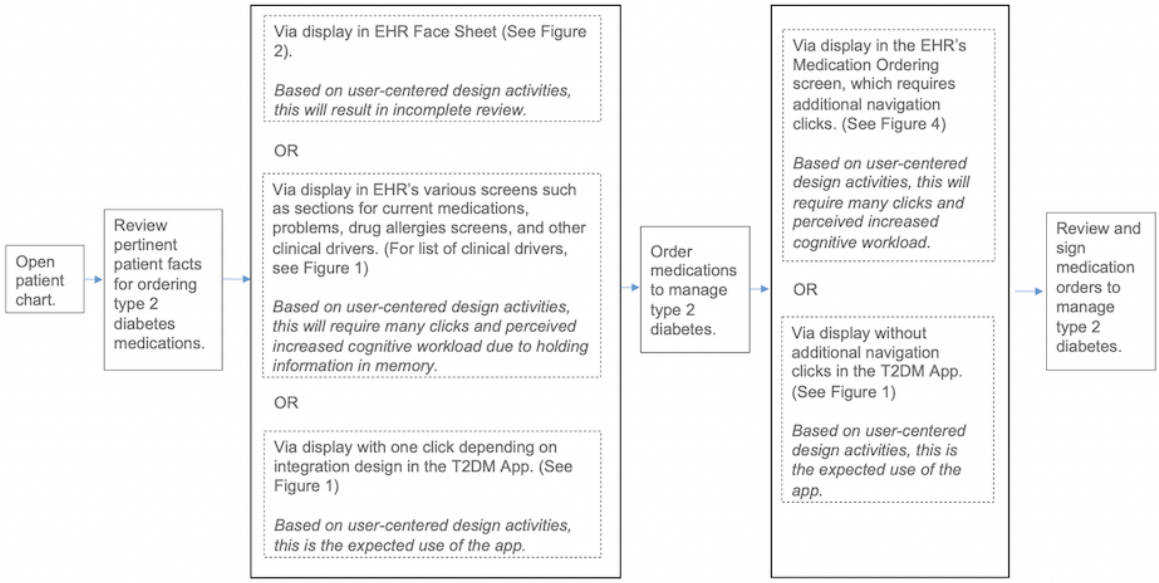
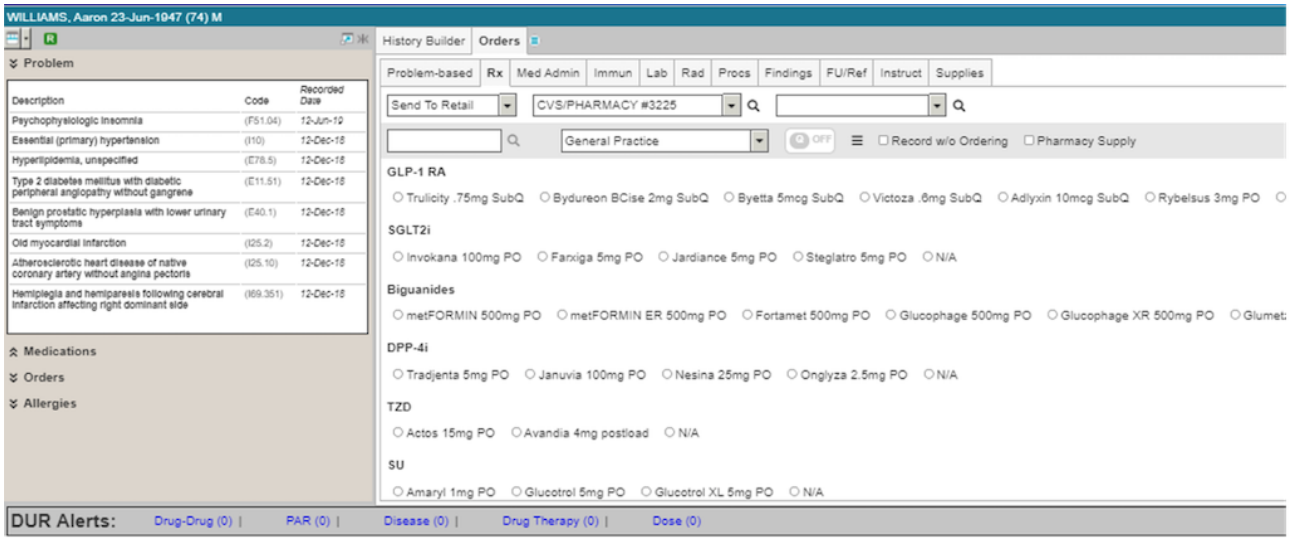


Figure 4. Mock medication ordering screen in the electronic health record. DPP-4i: dipeptidyl peptidase 4 inhibitor, GLP-1RA: glucagon-like peptide-1 receptor agonist, SGLT2i: sodium/glucose cotransporter-2 inhibitor, SU: sulfonylurea, TZD: thiazolidinedione.



Apparatus and Test Environment

The study was conducted between March 15 and 19, 2021, and because of continued COVID-19 social distancing restrictions, the test sessions were conducted remotely via a usability testing platform (Loop11) or via a web conference tool (Zoom or WebEx). One independent moderator was present with the participant throughout the session. Several sessions were observed by members of the product development team; two of the observing team members and the moderator asked participants clarifying questions during the session. The participants employed their own computers to display and interact with the apps. Although the environment with respect to seating, lighting, sound levels, temperature, and humidity varied, participants were most generally seated at a table or desk in a personal space within their home. In most cases, the participants shared their screens and interacted with the apps on their own desktops. In several cases, the moderator shared their screen and passed control of mouse and keyboard to the participant.

The method of data capture depended on the platform employed in the session. The Loop11 platform captured videos, task time, and screen review data; the web conference tools captured video only. In all cases, a notetaker recorded participant responses, and medication orders were collected manually from screen shots taken during the time of the study and determined by video review.

Participants also completed the Kano Model Survey. The Kano Model is a tool that can be used to prioritize the *critical to quality* characteristics, as defined by the *voice of the customer* [13]. The three categories identified by the model are as follows: (1) *Must-Have*: whatever the quality characteristic is, it must be present, such that if it is not, the customer will go elsewhere! (2) *Linear or Performance*: the better we are at meeting these needs, the happier the customer is. (3) *Exciter or Delighter*:

those qualities that the customer was not expecting but received as a bonus.

The task data collected included "pass" or "fail" (medication order adheres or does not adhere to evidence-based recommendations), task time, subjective comments from the Kano Model Survey, and video recordings of the computer screen and audio.

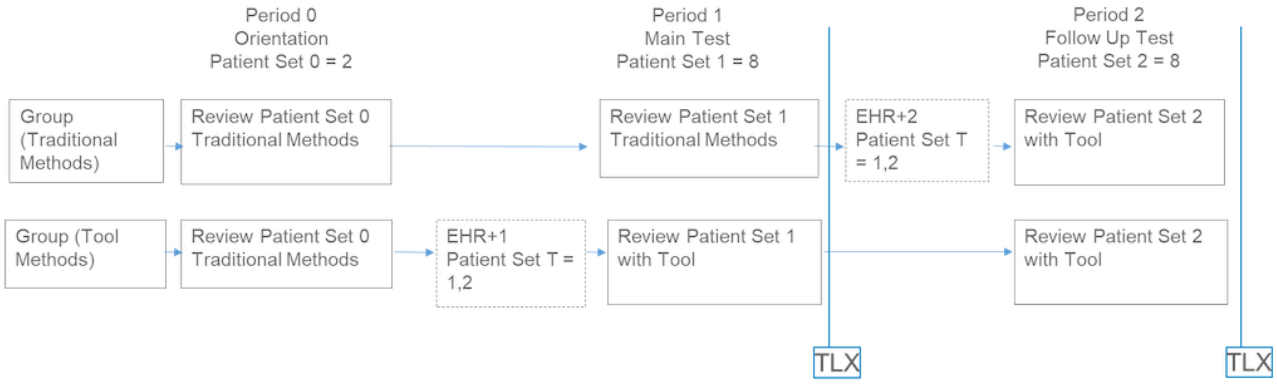
Study Design

The study employed a 2 (Group) × 2 (Time Period) mixed model simulated-use design (shown in Figure 5). Participants were randomly assigned to one of two groups (Traditional, Tool) with the constraint that each group was comprised of 50% MDs and 50% other medical licenses. Participants in both groups were asked to review two sets of eight mock patients and order diabetes medications. Participants in the Traditional group employed the mock EHR for reviewing and ordering in Period 1 and employed the prototype T2DM app for reviewing and ordering in Period 2. Participants in the Tool group employed the prototype T2DM app in both time periods. Both groups were afforded the opportunity to access any online resources that are typically used as part of their medication decision-making (eg, websites such as UpToDate, ADA evidence-based guidelines).

Each reviewing and ordering period was followed by administration of the NASA TLX to measure the perceived cognitive load associated with that performance period [14]. The NASA-TLX solicits ratings for mental demand, physical demand, temporal demand, performance, effort, and frustration on a scale from "Very Low" to "Very High." The web-based interface presented the scale as 10 radio buttons that were not numbered; responses were coded 0-100 consistent with extant literature [14]. Pairwise rankings of the indices were collected to generate weights for computation of the TLX score.

The data were analyzed to test the hypothesis that providers would more frequently align their orders with the ADA evidence-based guidelines when using the T2DM app, and that the process would be more efficient and less taxing. The hypothesis predicts performance advantages for the Tool group during the first period, and a greater change in duration for Traditional group in Period 2.

Figure 5. Schematic of the study design. EHR: electronic health record, TLX: Task Load Index.



Procedure

Each 2-hour session was conducted individually and began with guidance about the resolution of technical difficulties associated with the employed platforms and the informed consent process. The overall study process for both participant groups is shown in Figure 3.

Period 0

Orientation of participants to the EHR prototype. The moderator walked participants through the user interface and then participants completed 2 practice trials reviewing mock patient facts and ordering diabetes medications using traditional EHR methods. The purpose of the orientation was to ensure that participants understood the task. Both groups experienced the same protocol for this period.

Training Period

Orientation to the prototype T2DM app. The moderator walked participants through the user interface and then participants completed two practice trials reviewing mock patient facts and ordering diabetes medications using the prototype T2DM app. The Tool group experienced this protocol immediately before Period 1; the Traditional group experienced this protocol immediately before Period 2.

Period 1

Participants reviewed a set of 8 mock patients and ordered diabetes medications on the basis of each mock patient's facts. Participants were instructed to complete each order within 5 minutes and to do so as though they were at work. Participants in the Traditional group employed the mock EHR during this period; participants in the Tool group employed the prototype T2DM app.

TLX 1

Participants completed the NASA TLX survey aimed to measure perceived mental cognitive load for the tasks completed in Period 1.

Period 2

Participants reviewed another set of 8 mock patients and ordered diabetes medications on the basis of each mock patient's facts. As in period 1, participants were instructed to complete each order within 5 minutes and to do so as though they were at work. Participants in both groups employed the prototype T2DM app.

TLX 2

Participants completed the NASA TLX survey aimed to measure perceived cognitive load for the tasks completed in Period 2.

Data Analysis

We analyzed data using SPSS (version 25.0; IBM Corp). Dependent variables in the study (percent adherence to evidence-based guidelines, task time, screens reviewed, and perceived workload rating) were all subjected to a 2 × 2 mixed model analysis of variance (ANOVA) with Period (Period 1 vs Period 2) as the within-subjects factor and Group (Traditional vs Tool) as the between-subjects factor. Statistical significance was accepted at a level of $P<.05$.

Results

Prototype App Development

User Testing

Each round of user testing focused on different elements of the prototype app while also iteratively reviewing changes that resulted from the previous round or rounds. Changes became smaller and more focused with each successive round of testing until the final version was reached after round 8. Table 1 provides information about each round.

Table 1. Focus and results for each round of user testing.

Round	User testing focus	Changes informed by results
1	What clinical information prescribers used to inform medication ordering, medication ordering workflow, and visual presentation of information on the left side of the screen	Patient fact details, medication ordering workflow, and user interface improvements
2	What additional clinical information prescribers needed to inform diabetes medication ordering, details of the patient facts, and on and off guidelines for evidence-based prescribing	Patient facts details, order summary screen, and medication details relevant to prescribers
3	Adequacy of clinical information for prescribers to make appropriate medication decisions	Patient fact details, on and off guideline medication details, and order summary screen
4	Importance of various features for prescribers (eg, patient facts, clinical drivers, on and off guideline evidence–based medication table, and ordering)	Prioritization of included features
5	Ease of use for finding patient facts and ordering medication	User workflow and interface design
6	Ease of use for updated interface design	User workflow and interface design
7	Ease of use for finding patient facts and ordering or discontinuing medication	Ordering and discontinuing workflow
8	Utility and clarity of app user guide and product information	Presentation of information in user guide and product information

Subjective Comments

Comments made by user testing participants (Table 2) were used in determining design changes for each round.

Table 2. Sample comments from user testing.

Context	Comments
On laboratory results and interpretation	• “I’m the one with the medical degree, not the computer. I need to know where things are coming from.” [estimated glomerular filtration rate finding]
Presentation of clinical drivers	• “It’s amazing…I really like the way it pulls clinical drivers into one location so you can drive your recommendations based on that.” • “I'm not necessarily going to trust an app to be the end goal. If it has an explanation, I might have a little more trust.”
Presentation of medication cost	• “Cost should be specific to the patient’s insurance in order to be useful”
Flagging allergies	• “You need to know that [allergies] if you are looking at medications…. I would want that {allergies}to be more prominent.”
Drug utilization review checking in electronic medical records	• ”I didn’t realize it would take me to the EMR, I thought I would be able to do it through the app.”

SUS

The SUS scores were "average" or "above average" on each of the rounds of user testing where measured (Table 3). While the averages for rounds 1-4 were all above 68, the lower average for round 6 resulted from using an alternate design that participants perceived as less user-friendly; this design was subsequently abandoned in favor of the earlier design.

Table 3. System Usability Scale scores for 5 of the 8 rounds of user testing.

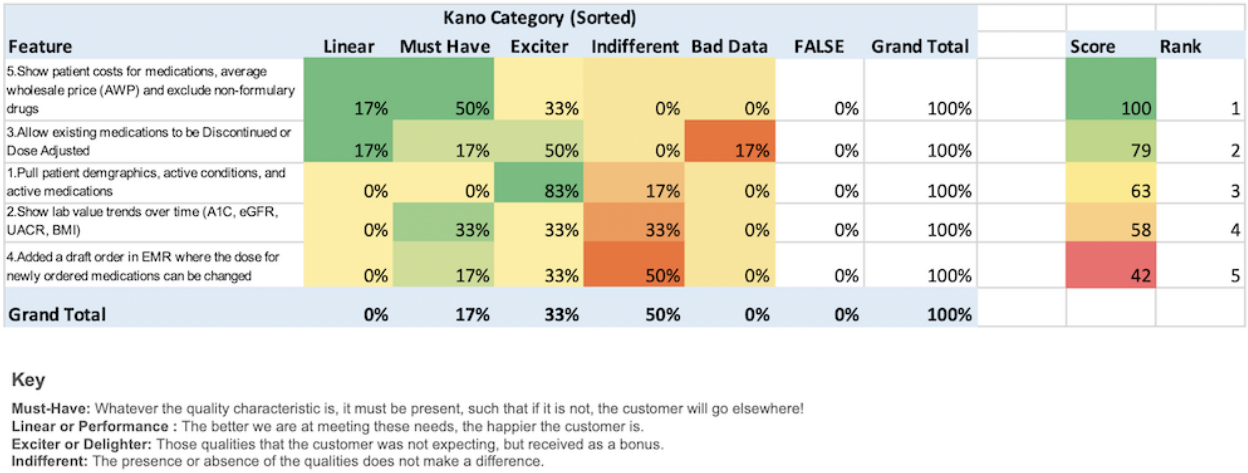
Round	Average score	Score range	Scores, n (individual scores)
1	83	60-95	7 (80, 87.5, 87.5, 87.5, 82.5, 60, 95)
2	73	65-78	6 (75, 65, 77.5, 72.5, 75, 75)
3	86	63-100	8 (95, 80, 62.5, 90, 97.5, 100, 85, 75)
4	81	63-95	6 (72.5, 87.5, 77.5, 95, 62.5, 90)
6	68	50-78	5 (50, 60, 90, 60, 77.5)

Validation Study

Combined Kano Model Study Results

In total, 14 of the validation study participants responded to the Kano Model Survey (Figure 6).

Figure 6. Kano Model Survey results. A1C: glycated hemoglobin, eGFR: estimated glomerular filtration rate, EMR: electronic medical record, UACR: urine albumin-creatinine ratio.

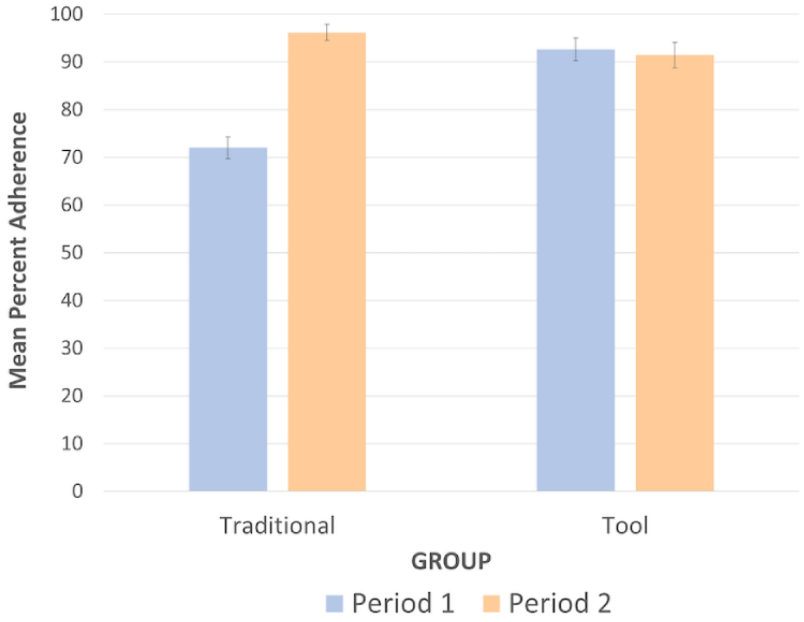


Alignment With Evidence-Based Guidelines

Orders were scored as *pass* if the participant ordered any medications that were on-guideline, discontinued any medications not aligned with guidelines (ie, a recommended discontinuation) or if the guidelines recommended against medication changes, and *fail* if any ordered medication was off-guideline.

The mean percentage adherence to ADA evidence-based guidelines is shown in Figure 7. There was a main effect of Group on adherence to guidelines ($F_{1,41}$=8.99, P=.005, =0.18). As hypothesized, medication ordering by the Tool group was more frequently aligned with guidelines (mean 92.1) than it was by the Traditional group (mean 84.2). Furthermore, there was a main effect of Period on adherence to guidelines ($F_{1,41}$=37.63, P<.001, =0.48). Medication ordering in Period 2 was more frequently aligned with guidelines (mean 93.7) than it was during Period 1 (mean 82.6).

Figure 7. Adherence to American Diabetes Association evidence-based guidelines.



However, there was a significant Group × Period interaction ($F_{1,41}$=45.43, P<.001, =0.53). For the Traditional group, ordering was more aligned with guidelines during Period 2, when the T2DM Tool was first introduced (mean 96.2), than it was in Period 1 (mean 72.0; t_{20}=–8.47, P<.001). For the Tool group, ordering was similarly aligned to guidelines during Period 1 (mean 92.6) and Period 2 (mean 91.5; t_{21}=0.46, P=.65).

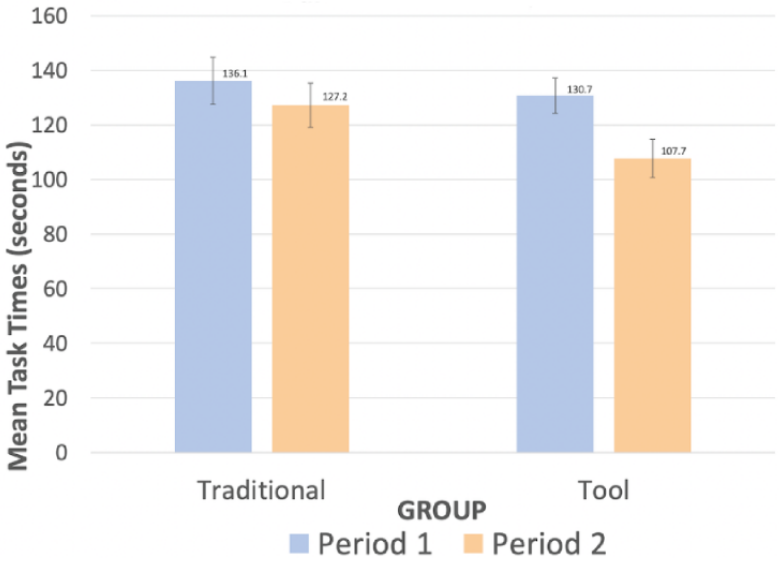
Of particular interest was a planned comparison between groups during Period 1. As hypothesized, medication ordering by the group using the prototype T2DM app was more frequently aligned to ADA evidence–based guidelines (mean 92.6) than medication ordering was by the group using the mock EHR (mean 72.0; t_{41}=–6.20, P<.001).

Task Time

Measurement of task times were impacted by technology issues (ie, poor internet connections), which resulted in extreme task times. Extreme task times were excluded based on a cutoff of 3 SD. Task time SDs were calculated using all available task time data. Task times greater than (and less than 3 SD) were excluded from the data analysis (ie, from the computation of each participant's mean for each period). The percentage of trials removed by the 3 SD trim was 1.8%. There were no task times 3 SD below the mean because that was a negative value.

The mean task time used for the completion of orders is shown in Figure 8. There was no main effect of Group on task time ($F_{1,41}$=1.46, P=.23, =0.03), but there was a main effect of Period ($F_{1,41}$=26.70, P<.001, =0.39). Medication ordering time was lower during Period 2 (mean 117.2) than it was in Period 1 (mean 133.4).

Figure 8. Task times during prescribing.



There was also a significant Group × Period interaction ($F_{1,41}$=5.23, P=.03,  =0.11). The Tool group completed prescription orders more quickly during the second period (mean 107.7) than the first period (mean 130.7; t_{21}=6.49, P<.001). For the Traditional group, task time did not significantly differ between Periods 1 (mean 136.1) and Period 2 (mean 127.2; t_{20}=1.74, P=.10). Finally, Period 1 task times also did not significantly differ between the two groups (t_{41}=0.50, P=.62).

Screens Reviewed During Prescribing

For technical reasons, screen review data was not captured for 9 participants from the Traditional group and 5 participants from the Tool group.

There was no main effect of Group on the number of screens reviewed ($F_{1,26}$=0.51, P=.48,  =0.02), and there was no main effect of Period ($F_{1,26}$=0.73, P=.40,  =0.03). Because the variance was found to be unequal across groups (by Box M and Levene tests), to reduce the probability of type 1 error, the α level was adjusted accordingly to be increasingly conservative (α'=.001). Based on the adjusted α level, there was no Group × Period interaction effect ($F_{1,26}$=7.16, P=.01,  =0.22).

Perceived Cognitive Load (NASA TLX)

Of the 43 participants, 9 participants from the Traditional group and 6 participants from the Tool group were not able to complete the survey for logistical or technical reasons.

There was no main effect of Group on the NASA TLX score ($F_{1,34}$=2.30, P=.14,  =0.06), and there was no main effect of Period ($F_{1,34}$=2.13, P=.15,  =0.06). There was no significant Group × Period interaction ($F_{1,34}$=0.23, P=.64,  =0.01). Finally, the perceived workload associated with Period 1 did not significantly differ between the two groups (t_{34}=1.14, P=.26).

Effect of Provider Type

As part of our analysis, we attempted to evaluate the effect of provider type on the studies using mixed ANOVAs with the clinical role (MD vs non-MD physicians) between subjects and test period (Period 1 and Period 2) within subjects. There was no significant main effect of clinical role on adherence to guidelines ($F_{1,41}$=4.05, P=.05,  =0.09), and role × time interaction ($F_{1,41}$=1.87, P=.18,  =0.04). Furthermore, there was no significant main effect of clinical role on test period ($F_{1,41}$=3.86, P=.06,  =0.09), and role × time interaction ($F_{1,41}$=.79, P=.38,  =0.02). Provider type analysis was not conducted for other parts of the study, partly owing to the sample size and initial findings.

Discussion

App Effectiveness, Efficiency, and User Satisfaction

This applied study aimed to develop a prototype medication recommender (T2DM) app via a thorough UCD process and evaluate the design for effectiveness, efficiency, and user satisfaction. The prototype T2DM app is considered *effective* if it supported care decisions that better aligned with evidence-based guidelines and more *efficient* as measured by reduced task time or reduced cognitive load for participants.

For the primary research question of whether medication ordering with the prototype T2DM app would be more frequently aligned to ADA evidence–based guidelines than medication ordering without the app and using the mock EHR (ie, app effectiveness), the prototype T2DM app proved effective when measured within as well as between participants. Providers who ordered medications using the mock EHR first became more aligned to ADA guidelines when they switched to the prototype T2DM app. When comparing the method employed in the first period, the group of providers that used the T2DM app were more aligned to ADA guidelines than the group that used the mock EHR.

For the primary research question of whether medication ordering with the prototype T2DM app would be accomplished more quickly than that without the app (with the mock EHR; ie, app efficiency): medication ordering was accomplished more quickly in the second period, which can reasonably be expected owing to practice. When comparing methods employed in the first period, the group of providers that used the prototype T2DM app did not complete their medication ordering more quickly than the group that used the mock EHR. However, the group that used the prototype T2DM app in both time periods improved more during the second period than the group that switched from the mock EHR to the prototype T2DM app. This outcome suggests that providers may be more efficient when using the T2DM app, but not until they have become more familiarized with its display and features. The greater efficiency in the second period for the tool-only group compared to the EHR group is likely owing to the inherent design and features of the prototype app: it is fit-for-purpose, avoids navigation distractions in the chart, and pulls information from different parts of the EHR to prefill the clinical facts. There was also no significant effect of the difference in provider type in alignment to guidelines or task times.

For the primary research question of whether medication ordering with the prototype T2DM app would be accomplished with lower perceived cognitive load than that without the app (with the mock EHR; ie, app efficiency), NASA TLX scores were generally low. Although benchmarking is not possible with the TLX, a published analysis of 237 studies would place our overall mean of 42.5 at approximately the 35th percentile [15]. Clearly, medication ordering was not exceptionally taxing with either interface, but the failure to find differences should not be interpreted as a floor effect. Despite changes in ordering behavior, the perceived cognitive load was relatively stable.

Finally, the prototype T2DM app was evaluated for *user satisfaction* throughout the iterative rounds of user testing. SUS scores throughout the iterative design process were at or above 68, which is considered average across industries. Round 6, the last iterative user testing during which the SUS was administered, had an average SUS score of 68. This was the lowest average SUS score obtained during the design process. The design evaluated in Round 6 went against several good user interface principles, and the prototype was not as interactive as the previous prototypes; hence, participants were not able to experience many of the previously identified valuable features. SUS scores for EHR systems have been shown to be lower than the average (68) from across different industries [12]. Melnick et al [16] benchmarked EHR usability by having 870 physicians complete the SUS questionnaire on the basis of their experiences with their own EHR system. The mean SUS score was 45, a score characterized as not acceptable and given a grade of F. Thus, although we caution against making direct comparisons, these data provide a favorable background for considering the usability of the prototype T2DM app.

The information from this prototype testing will contribute to the development of a real CDS-T2DM app, which will leverage SMART (Substitutable Medical Applications and Reusable Technologies) on FHIR (Fast Healthcare Interoperability Resource) technology. It is important to note that clinical decision support apps may be subject to FDA (Food and Drug Administration) review and approval. The scope of FDA oversight and compliance with any related regulatory requirements was beyond the scope of this study on the prototype app.

Limitations

We acknowledge some limitations of our study, including (1) a small sample size, (2) participant attrition in certain rounds of the study and NASA TLX assessment, and (3) difference in platforms for observation (The Loop11 platform captured videos, task time and screen review data; the web conference tools captured video only). Furthermore, owing to the limited sample size, we were neither able to reliably assess the effect of provider type (MD vs non-MD physicians) across the studies nor the impact of speed to adoption on task time. We believe that additional research is needed to further evaluate the effectiveness and efficiency of the T2DM app as it launches and becomes more widely used.

Conclusions

T2DM and prediabetes affect millions of Americans, often resulting in harmful complications, additional chronic conditions, disability, premature death, and significant patient and system costs. Adherence to clinical treatment guidelines can improve patient health outcomes and reduce patient and system costs. Complexity of medical guidelines combined with limitations on providers' time can impede guideline adherence. While CDSSs can help, lack of user involvement in their development can further impede progress and result in mistrust in technology solutions.

Through deployment of the UCD process, we developed a prototype of a medication recommender app that promises to improve adherence to ADA evidence–based guidelines and support a more efficient and user-friendly ordering process for the provider in the management of T2DM. CDSSs offer promising solutions for closing the gap between provider behavior and evidence-based practice, and this study suggests that realizing their full promise may depend on greater attention to design from a user-centered perspective. Such a process could be beneficial in developing effective CDSSs for other conditions and tasks with associated improvement in quality and costs for patients, providers, and the health care system.

Acknowledgments

The authors would like to thank the team at Userview Inc led by Janey Barnes who helped with study operations and analysis and Jessica Knight at OptumLabs, who helped review the manuscript and guided the team through the internal medical and legal review.

Authors' Contributions

KL, DK, HH, and RE were involved in the design of the study. RE and PM ran the usability studies and secured the study operations and analytics team. SC supported the usability studies. QH contributed to building the app and using feedback from study to refine app development. BA, KL, and RE helped with the interpretation of findings from the studies and analyses. BA, RE, PM, DK, KL, and BK co-wrote an early draft of the paper. All authors reviewed and edited the subsequent drafts. All authors approved the final version of the manuscript.

Conflicts of Interest

At the time of submission, all authors but BK are full or contract employees of OptumHealth, a part of Optum. BK is an employee of OptumLabs. All authors but PM own Optum/UHG stocks.

References

1. National Diabetes Statistics Report, 2020. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/diabetes/data/statistics-report/index.html [accessed 2021-07-21]
2. Economic Impact of and Treatment Options for Type 2 Diabetes. AJMC. 2017 Aug 21. URL: https://www.ajmc.com/view/economic-impact-of-and-treatment-options-for-type-2-diabetes-article [accessed 2021-07-21]

3. American Diabetes Association. 6. Glycemic Targets:. Diabetes Care 2021 Jan;44(Suppl 1):S73-S84. [doi: 10.2337/dc21-S006] [Medline: 33298417]
4. Curtis HJ, Dennis JM, Shields BM, Walker AJ, Bacon S, Hattersley AT, et al. Time trends and geographical variation in prescribing of drugs for diabetes in England from 1998 to 2017. Diabetes Obes Metab 2018 Sep;20(9):2159-2168 [FREE Full text] [doi: 10.1111/dom.13346] [Medline: 29732725]
5. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17 [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
6. Ali SM, Giordano R, Lakhani S, Walker DM. A review of randomized controlled trials of medical record powered clinical decision support system to improve quality of diabetes care. Int J Med Inform 2016 Mar;87:91-100. [doi: 10.1016/j.ijmedinf.2015.12.017] [Medline: 26806716]
7. O'Connor PJ, Sperl-Hillen JM, Rush WA, Johnson PE, Amundson GH, Asche SE, et al. Impact of electronic health record clinical decision support on diabetes care: a randomized trial. Ann Fam Med 2011;9(1):12-21 [FREE Full text] [doi: 10.1370/afm.1196] [Medline: 21242556]
8. Holbrook A, Thabane L, Keshavjee K, Dolovich L, Bernstein B, Chan D, COMPETE II Investigators. Individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. CMAJ 2009 Jul 07;181(1-2):37-44 [FREE Full text] [doi: 10.1503/cmaj.081272] [Medline: 19581618]
9. Meigs JB, Cagliero E, Dubey A, Murphy-Sheehy P, Gildesgame C, Chueh H, et al. A controlled trial of web-based diabetes disease management: the MGH diabetes primary care improvement project. Diabetes Care 2003 Mar;26(3):750-757. [doi: 10.2337/diacare.26.3.750] [Medline: 12610033]
10. Sim LLW, Ban KHK, Tan TW, Sethi SK, Loh TP. Development of a clinical decision support system for diabetes care: A pilot study. PLoS One 2017;12(2):e0173021 [FREE Full text] [doi: 10.1371/journal.pone.0173021] [Medline: 28235017]
11. Introduction: Standards of Medical Care in Diabetes-2020. Diabetes Care 2020 Jan;43(Suppl 1):S1-S2. [doi: 10.2337/dc20-Sint] [Medline: 31862741]
12. Measuring Usability with the System Usability Scale (SUS). MeasuringU. 2011 Feb 03. URL: https://measuringu.com/sus/ [accessed 2021-07-21]
13. Spool JM. Understanding the Kano Model – A Tool for Sophisticated Designers. Center Centre Inc. 2018 Nov 06. URL: https://articles.uie.com/kano_model/ [accessed 2021-08-11]
14. Hart S, Staveland L. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Adv Psychol 1988;52:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
15. Grier RA. How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores. Proc Hum Factors Ergon Soc Annu Meet 2016 Dec 20;59(1):1727-1731. [doi: 10.1177/1541931215591373]
16. Melnick ER, Dyrbye LN, Sinsky CA, Trockel M, West CP, Nedelec L, et al. The Association Between Perceived Electronic Health Record Usability and Professional Burnout Among US Physicians. Mayo Clin Proc 2020 Mar;95(3):476-487 [FREE Full text] [doi: 10.1016/j.mayocp.2019.09.024] [Medline: 31735343]

Abbreviations

ADA: American Diabetes Association
ANOVA: analysis of variance
CDS: clinical decision support
CDSS: clinical decision support system
EHR: electronic health record
FDA: Food and Drug Administration
FHIR: Fast Healthcare Interoperability Resource
HbA_{1c}: glycated hemoglobin
NASA TLX: National Aeronautics and Space Administration Task Load Index
SMART: Substitutable Medical Applications and Reusable Technologies
SUS: System Usability Scale
T2DM: type 2 diabetes mellitus
UCD: user-centered design

Edited by G Eysenbach; submitted 14.09.21; peer-reviewed by J Walsh, R Marshall, P Macharia; comments to author 06.10.21; revised version received 03.11.21; accepted 12.11.21; published 18.01.22.

Please cite as:
Larsen K, Akindele B, Head H, Evans R, Mehta P, Hlatky Q, Krause B, Chen S, King D
Developing a User-Centered Digital Clinical Decision Support App for Evidence-Based Medication Recommendations for Type 2 Diabetes Mellitus: Prototype User Testing and Validation Study
JMIR Hum Factors 2022;9(1):e33470
URL: https://humanfactors.jmir.org/2022/1/e33470
doi:10.2196/33470
PMID:34784293

Original Paper

Improving Pelvic Floor Muscle Training Adherence Among Pregnant Women: Validation Study

Aida Jaffar[1,2], MMed; Sherina Mohd-Sidik[1], PhD; Chai Nien Foo[3], PhD; Novia Admodisastro[4], PhD; Sobihatun Nur Abdul Salam[5], PhD; Noor Diana Ismail[6], MMed

[1]Department of Psychiatry, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Selangor, Malaysia

[2]Primary Care Unit, Faculty of Medicine and Defence Health, Universiti Pertahanan Nasional Malaysia, Wilayah Persekutuan, Malaysia

[3]Department of Population Medicine, Universiti Tunku Abdul Rahman, Selangor, Malaysia

[4]Software Engineering & Information System Department, Faculty of Computer Science & Information Technology, Universiti Putra Malaysia, Selangor, Malaysia

[5]School of Multimedia Technology and Communication, College of Arts and Sciences, Universiti Utara Malaysia, Kedah, Malaysia

[6]Klinik Kesihatan Bt 9 Cheras, Ministry of Health, Selangor, Malaysia

Corresponding Author:
Sherina Mohd-Sidik, PhD
Department of Psychiatry
Faculty of Medicine and Health Sciences
Universiti Putra Malaysia
Serdang
Selangor, 43400
Malaysia
Phone: 60 3 9769 2541
Fax: 60 3 9769 2706
Email: sherina@upm.edu.my

Related Article:

This is a corrected version. See correction statement: https://humanfactors.jmir.org/2022/2/e38175

Abstract

Background: Mobile health apps, for example, the Tät, have been shown to be potentially effective in improving pelvic floor muscle training (PFMT) among women, but they have not yet been studied among pregnant women. Adherence to daily PFMT will improve pelvic floor muscle strength leading to urinary incontinence (UI) improvement during the pregnancy.

Objective: This study aims to document the validation process in developing the Kegel Exercise Pregnancy Training app, which was designed to improve the PFMT adherence among pregnant women.

Methods: We utilized an intervention mapping approach incorporated within the mobile health development and evaluation framework. The framework involved the following steps: (1) conceptualization, (2) formative research, (3) pretesting, (4) pilot testing, (5) randomized controlled trial, and (6) qualitative research. The user-centered design-11 checklist was used to evaluate the user-centeredness properties of the app.

Results: A cross-sectional study was conducted to better understand PFMT and UI among 440 pregnant women. The study reported a UI prevalence of 40.9% (180/440), with less than half having good PFMT practice despite their good knowledge. Five focus group discussions were conducted to understand the app design preferred by pregnant women. They agreed a more straightforward design should be used for better app usability. From these findings, a prototype was designed and developed accordingly, and the process conformed to the user-centered design–11 (UCD-11) checklist. A PFMT app was developed based on the mHealth development and evaluation framework model, emphasizing higher user involvement in the application design and development. The application was expected to improve its usability, acceptability, and ease of use.

Conclusions: The Kegel Exercise Pregnancy Training app was validated using a thorough design and development process to ensure its effectiveness in evaluating the usability of the final prototype in our future randomized control trial study.

(JMIR Hum Factors 2022;9(1):e30989) doi:10.2196/30989

KEYWORDS

User-centered design; mHealth app; Digital intervention; mHealth Development and Evaluation Framework; Usability; Acceptability; Pelvic Floor Muscle Training; Urinary incontinence; Pregnancy

Introduction

Background

Urinary incontinence (UI) is defined as an involuntary urinary leakage [1] and has been affecting up to 57.7% of women in low-income countries [2]. Pregnancy affects hormonal changes, which may add an additional risk of UI among adult women [3]. Additionally, other factors, such as damage to the periurethral, paraurethral, and pubo-urethral connective tissues, may occur during pregnancy, labor and delivery, or with obesity, which may impact the urethra and bladder neck position at rest, leading to UI [4]. Another recent meta-analysis, which included studies from 1998 to 2018, stated that about half of pregnant women (41.0%) experienced UI, and it negatively affects their quality of life (QoL) [5].

Women with UI had an unpredictable demand to use the bathroom and became more aware of the bathroom location when shopping and traveling. They felt fear and easily frustrated when they were unable to get to the bathroom on time [6]. During the nighttime, they visited the bathroom a few times, and their sleep was affected significantly. Interrupted sleep led to problems in occupational functioning, as well as psychological functioning, as it worsened depression [7]. Pregnant women may experience a multidimensional negative impact on QoL, such as social-emotional relationships, physical activities, employment issues, limitation to travel, sleep disruption, and obstacles in performing their prayers [8]. Pregnant women who experience UI during pregnancy reported having 5 times the risk of having UI during their postpartum period [9]. Therefore, there is an urgent need to screen and treat them as early as possible by performing pelvic floor muscle training (PFMT).

PFMT or Kegel exercise is an essential exercise among pregnant women. Strengthening their pelvic floor muscles is recommended as it is minimally invasive and does not involve any complications [1,10]. The advantages of the exercise are to shorten the duration of the second stage of labor, reduce severe perineal lacerations [11], and shorten the painful experience of the postpartum period. Pregnant women should be aware of the benefit and be able to perform the correct techniques of the exercise.

Regretfully, not all pregnant women are aware of PFMT despite UI affecting their daily activities because PFMT uptake may be constrained by the antenatal service provision or challenges in accessing services at a primary care clinic [12]. Pregnant women experience challenges related to PFMT education as it involves an individualized approach, for example, the specific language, different levels of health literacy, and cultural variations [13]. In some cultures, the anatomical involvement may be sensitive to certain women [14], leading to necessary adjustments in disseminating the correct information of PFMT according to those cultures.

Additionally, there may be limited attention toward pelvic floor health during pregnancy provided by health care providers (HCPs) as it may be unclear to whom these professional responsibilities belong [13]. PFMT is actually under the responsibility of a physiotherapist; however, pregnant women are under antenatal care follow-up, which is conducted mainly in primary care clinics where there is limited access to physiotherapists [13]. Moreover, the lack of standardized guidelines, inadequate information, and a lack of continuity of care may result in organizational variation in antenatal care services, which worsens the accessibility and acceptability of PFMT services [13].

Regarding individual factors causing training barriers, only one-tenth of pregnant women seek help due to the misperception that UI will resolve by itself [15], assuming that it is “normal” to have UI, and having a misconception that pelvic training will lead to miscarriage [16]. Hence, women face difficulties in achieving the necessary knowledge and skills, resulting in poor attitude and adherence towards pelvic training [13]. Adherence to daily training is one of the most important prognostic factors for PFMT effectiveness in both the short term and longer term [17]. A new method to disseminate PFMT education is necessary to manage both (individual and HCPs) barriers.

Mobile Health App

Mobile health (mHealth) is defined as “a medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices” [18]. The National Institute for Health and Care Excellence Guideline has categorized the mHealth apps into three-tier functional classifications:

1. Tier A: An app that provides health and social care services without measurable patient outcomes.
2. Tier B: An app that provides health and lifestyle information, health monitoring, or patient-health care professional communication.
3. Tier C: An app with interventions [19].

With regards to the tier C classification, the mHealth app intervention consists of six items which are (1) addressing preventative behavior change, such as smoking; (2) addressing self-management specific behaviors using behavior change techniques; (3) guiding the treatment; (4) providing active monitoring, for example, tracking patients’ location; (5) providing diagnosis, care, or calculated treatment; and (6) providing or guiding the diagnoses [19].

What makes mHealth a powerful tool for behavior modification are its strengths, which include its ease of access and user-friendliness, resulting in its widespread adoption worldwide [20,21]. mHealth can be used by an individual anywhere, at any

time, and patients can communicate with health care providers or even a chatbot on specific issues related to the app they are using [22,23]. The strength of a well-designed app is the ability to be well-accepted by the users, change their attitudes, and reduce the acceptability barriers in receiving health care services.

However, qualitative reviews on midwives reported unfavorable findings which did not fully support the apps for several reasons [24]. Midwives were concerned about the accuracy of the apps [25] and their negative impacts on the patient-professional relationship, such as shifting the patient's trust from trusted to untrusted sources [24]. Hence, there is a need to design a validated pregnancy mHealth app that has undergone the necessary steps based on a framework to improve its effectiveness.

User-Centered Design

An mHealth app needs to be carefully designed to ensure its impact on the users and its effectiveness. Despite mHealth becoming popular, there is still limited evidence of its effectiveness [26], which is most probably due to the unmet need of incorporating users in the design process [27]. mHealth apps designed *for* the users but not *with* the users will lead to high rates of technology rejection [28,29]. Hence, mHealth apps should be designed to fulfill the user's requirement using a user-centered design (UCD) framework.

Designing and developing an mHealth app using the user-centered element by involving the users in all stages has its robustness and is a gold standard approach in accomplishing mHealth apps that are useful, easy to use, and satisfying to the users [30]. For example, an mHealth app on the diet and oral health for parents of preschool children, which was developed based on their needs, included information on how to prevent oral disease in their children and has been scored with good usability [31]. UCD requires an iterative design process to understand and internalize the users' needs, goals, strengths, limitations, contexts, and intuitive processes [32]. Additionally, the iterative process includes observing the users' interaction with the app during the development process [32]. After understanding the importance and iterative UCD element, the app is then developed and incorporated within the software development life cycle (SDLC), which has been reported as the most liked technique to develop a high-quality software product [33].

The mHealth development and evaluation framework was developed using the iteration SDLC version, and it includes six stages: (1) conceptualization, (2) formative research, (3) pretesting, (4) pilot testing, (5) randomized controlled trial, and (6) qualitative research for further refinement before moving to a more scaled-up intervention [34].

Conceptualization involved experts' decisions regarding the theoretical basis, reviewing the evidence, and planning the development process via several brainstorming sessions. The research team includes a persuasion element to improve the user's engagement with the app.

Persuasive systems may be defined as "computerized software or information systems designed to reinforce, change, or shape attitudes or behaviors or both without using coercion or deception" [35]. the persuasive system design (PSD) categorized the 28 strategies into four main categories: primary task support, dialogue support, credibility support, and social support [35]. Among the most PSD used was the primary task support using the self-monitoring tracking [36] in physical activity apps.

Formative research, which is the next stage, involved focus group discussions (FGDs) with pregnant women to determine how they used the app and which design they preferred to use on their mobile phones. The pretesting stage, which stressed the importance of the message context (PFMT adherence), was set and strengthened the collection of responses from pregnant women. The reason for the differences was that not all participants would be able to grasp every message, and key messages (PFMT) could be repeated in different contexts to reach more pregnant women.

A pilot study stage was conducted to obtain further feedback from pregnant women after they had used it over several weeks. The next stage is the randomized control trial stage aimed at obtaining rigorous evidence and, finally, a qualitative study to explore the use of the app in depth.

This study aims to document the validation process of a newly designed mHealth app called the Kegel Exercise Pregnancy Training (KEPT) app running on the Android platform. The KEPT app is intended to deliver training sessions, send reminders, and chart PFMT and UI symptoms. The expectation of the app is to deliver the correct method of PFMT efficiently and conveniently according to the pregnant women's time and place, without the need to be in the clinic or consult with physiotherapists or doctors. The KEPT app is expected to fill the information gap between clinical visits and has undergone its usability evaluation by the experts [37].

Methods

Overview

We incorporated an intervention mapping (IM) approach with a UCD SDLC framework called the mHealth development and evaluation framework. This study focused on conceptualizing the app and using the UCD-11 checklist to evaluate the user-centeredness properties of the app.

Intervention Mapping

IM is a framework that was designed to plan an effective intervention from the needs assessment up to its evaluation. It comprises stepwise decision-making in developing, implementing, and evaluating interventions using community-based research methods [38,39]. According to Fernandez et al, the participation from the community is to ensure that the intervention matches priority population needs and intervention contexts [39]. The characteristics of this approach involve three aspects that are applied during the intervention planning process: (1) participatory planning, (2) the comprehensive theory used, and (3) an ecological and systems approach for understanding health problems and intervening to address them [40].

Intervention development, according to IM, comprises six steps: (1) demonstrate a comprehensive understanding of the health

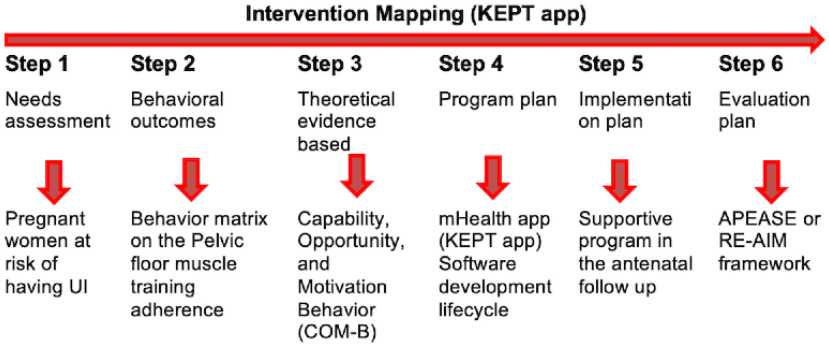
problem; (2) outline the behavioral and environmental outcomes; (3) identify theory-based and evidence-based behavior change methods that affect the determinants and translate these into practical applications that fit the intervention context; (4) combine the intervention components into a coherent program that uses delivery channels that fit the context; (5) develop implementation strategies to facilitate adoption, implementation, and maintenance of the program; and (6) plan both process and outcome evaluations to assess program implementation, and efficacy or effectiveness [39].

Accordingly, the intervention mapping of this mHealth app involved the stepwise approach from the needs assessment followed with other steps as illustrated in Figure 1.

The behavior matrix of this intervention can be divided into knowledge (14 questions), attitude (8 questions), and practices (4 questions) with regards to pelvic floor muscle exercise (PFME). The outcomes of the intention are self-efficacy (17 questions) [41] and adherence (6 questions) [42]. A few examples are listed in Textbox 1. Additional information about the questionnaire has been presented in the results section and published elsewhere [43].

This study involved documenting the project identification stage followed by the user experience design (Figure 2).

Figure 1. Intervention mapping framework of the KEPT app development. APEASE: affordability, practicability, effectiveness and cost-effectiveness, acceptability, side-effects and safety, equity; KEPT: Kegel Exercise Pregnancy Training; RE-AIM: reach, effectiveness, adoption, implementation, maintenance; UI: urinary incontinence.



Textbox 1. Behavior matrix of the KEPT app intervention to assess the determinants of pregnant women's adherence towards pelvic floor muscle training. KEPT: Kegel Exercise Pregnancy Training; PFME: pelvic floor exercise; UI: urinary incontinence.

Knowledge (K):

- Pelvic floor muscles involvement in pelvic exercise.
- Benefits of pelvic exercise.
- Methods in performing the pelvic exercise.

Attitude (A):

- I should practice PFME to prevent or treat UI.
- I should practice PFME to prevent uterine prolapsed.
- I feel that PFME is boring.

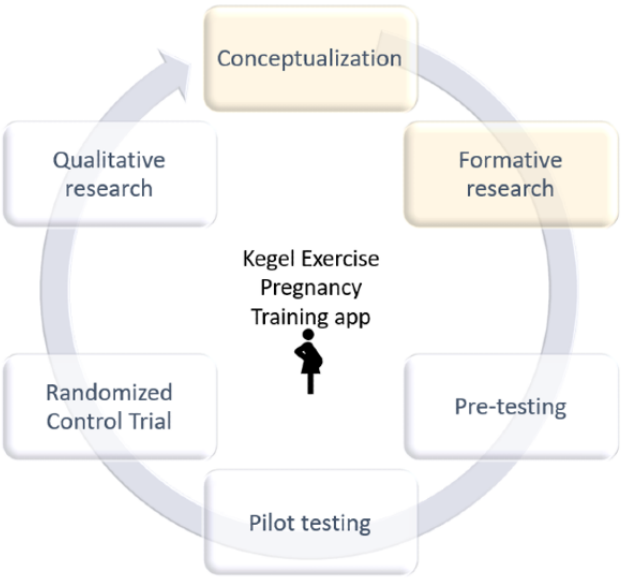
Practice (P):

- I had performed PFME when I was not pregnant.
- I have spent time performing PFME.
- I have tried to search for information regarding PFME.

Outcomes:

- Self-Efficacy (SE; how confident you can):
  - Perform pelvic exercises on your own.
  - Remember to perform exercises every day.
  - Perform the exercises at least three times a week.
- Adherence (AD):
  - I do my exercises as often as recommended.
  - I forget to do my exercises.
  - I do fewer exercises than recommended by my health care professional.

Figure 2. KEPT app development and evaluation framework. KEPT: Kegel Exercise Pregnancy Training.



In the project identification phase (conceptualization), we determined pregnant women's (users) understanding of pelvic exercise and the severity of this problem (UI) affecting their QoL. From this foundation, we produced a low-fidelity design, and we conducted FGDs to find out the users' experience using the design.

Conceptualization

The team assessed the latest evidence regarding the users' UI and pelvic floor training, finalizing the relevant theoretical behavioral change to be used and planning the development process according to the IM.

Cross-Sectional Study

A detailed understanding of at-risk pregnant women was established by conducting a cross-sectional study at a primary care clinic [43]. Within this study, we also determined the prevalence and severity of UI and its impact on participants' QoL at a primary care clinic (user's study setting) [43,44]. The findings from this study provided input for the content of their educational videos and short notes on PFMT [45,46], which were captured as frequently asked questions (FAQ).

Behavioral Change Theory

The team brainstormed and decided to identify theory-based and evidence-based behavior change methods concerning PFMT and UI. The interventions were found to be effective when they were developed based on behavioral changed theories, for example, the social cognitive theory and the health belief model [47]. Another overarching framework of behavior used to identify appropriate targets for enhancing adherence in clinical practice is the capability, opportunity, and motivation-behavior (COM-B) model [48,49].

The COM-B model states that motivation is a crucial source of strength to perform a certain behavior, with the assistance from the capability (physical and psychological) and opportunity (physical and social) to engage in the behavior [50], and the strength of motivation to engage in the behavior must be greater than for any other competing behavior.

Formative Research

A focus group discussion (FGD) study was conducted to understand which design was preferred by the end users. A low-fidelity prototype was given to 5 groups of end users as they chose their preferred design, including their reasoning.

Ethics Approval

This study was conducted according to the guidelines of the Declaration of Helsinki. The study approvals were obtained in August 2019 from the Ethics Committee for Research Involving Human Subjects, Universiti Putra Malaysia (JKEUPM-2019-368) and the Medical Research and Ethics Committee (MREC), Ministry of Health Malaysia (NMRR-19-412-45606).

Results

Conceptualization (Needs Assessment)

Cross-Sectional Study

A cross-sectional study was conducted among 440 pregnant women in a semi-urban primary care clinic with a response rate of 72.1% (440/610). The validated study instruments used consisted of socio-demography, knowledge, attitude, and practice on PFME and the International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form to determine UI among the respondents. The study reported that good knowledge, attitude, and practice scores towards pelvic exercise among pregnant women were 58.0% (255/440), 46.6% (205/440), and 45.2% (199/440), respectively, with further details published elsewhere [43]. The result indicated that pregnant women were not exercising despite having good knowledge (Table 1).

Table 1. Socio-demographic data among pregnant women (N=440).

Socio-demographic	Value
Age in years, mean (SD)[a]	29.84 (4.69)
Ethnicity, n (%)	
Malay	356(80.9)
Chinese	41 (9.3)
Indian	29 (6.6)
Others	14 (3.2)
Financial status, n (%)[b]	
Less than RM3000	128 (34.5)
RM3000-RM6274	176 (47.4)
RM6275-RM13147	59 (15.9)
RM13148 and above	8 (2.2)
BMI, n (%)	
Underweight	43 (9.8)
Normal	143 (32.5)
Overweight	144 (32.7)
Obese	110 (25.0)
History of Cesarean Section, n (%)	
No	181 (69.9)
Yes	78 (30.1)
Previous history UI[c], n (%)	
No	354 (80.5)
Yes	85 (19.3)
Pregnancy category, n (%)	
Primigravida	170 (38.6)
Multigravida	230 (52.3)
Grand multigravida	40 (9.1)
Trimester, n (%)	
First trimester	53 (12.0)
Second trimester	152 (34.5)
Third trimester	235 (53.4)
Category of UI, n (%)	
No UI	260 (59.1)
Slight UI	95 (21.6)
Moderate UI	80 (18.2)
Severe UI	5 (1.1)
Knowledge on PFME[d], n (%)	
Poor	185 (42.0)
Good	255 (58.0)
Attitude towards PFME, n (%)	
Poor	235 (53.4)
Good	205 (46.6)

Socio-demographic	Value
Practice on PFME, n (%)	
Poor	241 (54.8)
Good	199 (45.2)

[a]The median age is 30.0 years (27-33).

[b]A currency exchange rate of 1MYR = US $0.24 is applicable. The median income is RM4000 (2000-6000).

[c]UI: urinary incontinence.

[d]PFME: pelvic floor muscle exercise.

Pregnant women were unaware that the pelvic floor muscles were involved in controlling the anus and vagina. Some of them (184/440, 41.8%) were unaware of the correct duration and frequency of the exercise (158/440, 35.9%; Table 2). Hence, pregnant women were unaware of the anatomy of the pelvic floor muscles and the correct techniques of the pelvic exercise.

Only a quarter of them (111/440, 25.2%) strongly agreed that they should be taught the exercise and less than a fifth of them (83/440, 18.9%) agreed to put any effort into doing the exercise and practicing it. Pregnant women's attitude toward PFME in this study was less favorable (Table 3).

Very few of them (15/440, 3.4%) were performing the exercise whether or not they were pregnant, and only a few of them (12/440, 12.7%) have been practicing the exercise regularly (Table 4).

Based on this cross-sectional study, there was a need to have the PFMT intervention in delivering the correct pelvic exercise knowledge and self-efficacy improvement among pregnant women in order to improve their pelvic floor muscle strength.

Table 2. Pregnant women's responses about their knowledge of pelvic floor muscle exercises.

Knowledge on PFME[a]	Correct, n (%)
PFME muscles are situated in the pubic region	228 (51.8)
PFME involves muscles in the anal region	196 (44.5)
Vagina muscles are not involved in PFME	49 (11.1)
PFM[b] are important in controlling bladder function	296 (67.3)
PFM is not involved in controlling the anus	80 (18.2)
PFM is not involved in tightening the vagina	102 (23.2)
PFME can tighten buttocks muscles	193 (43.9)
PFME can prevent UI[c] during laughing/sneezing/weight bearing	292 (66.4)
PFME can prevent/treat uterine prolapse	244 (55.5)
PFME can be done at any time	315 (71.6)
PFME can be done while performing daily activities	248 (56.4)
Muscles involved should be contracted for 8 seconds	184 (41.8)
PFM should be contracted 8-10 times per exercise	170 (38.6)
PFME should be done at least 3x a day (morning, afternoon, and night)	158 (35.9)

[a]PFME: pelvic floor muscle exercise.

[b]PFM: pelvic floor muscle.

[c]UI: urinary incontinence.

Table 3. Pregnant women's attitude towards pelvic floor muscle exercise.

Attitude on PFME[a]	Strongly agree, n (%)
PFME should be done by all women	45 (10.2)
I should practice PFME to prevent/treat UI[b]	86 (19.5)
I should practice PFME to prevent uterine prolapse	72 (16.4)
I feel that PFME is boring	3 (0.7)
PFME should be taught to all antenatal mothers at antenatal clinics	111 (25.2)
I support those who want to perform PFME	113 (25.7)
I view that PFME can increase sexual satisfaction	73 (16.6)
I will put in the effort to search for info about PFME	83 (18.9)

[a]PFME: pelvic floor exercise.

[b]UI: urinary incontinence.

Table 4. Pregnant women's practice behavior towards pelvic floor muscle exercises.

PFME[a] practices	Always, n (%)
I have performed PFME when not pregnant	15 (3.4)
I have spent time performing PFME	12 (12.7)
I have discussed PFME with friends	7 (1.6)
I have tried to search for info about PFME	12 (2.7)

[a]PFME: pelvic floor muscle exercise.

Conceptualization (Theoretical Framework)

The theoretical framework of choice was the COM-B model with a combination of the PSD. The PSD is used with the COM-B is to reinforce the behavior voluntarily and to shape the attitude towards PFMT.

The COM-B model that was built into this app was expected to motivate pregnant women to perform PFMT regularly. Capability was divided into two subcategories: (1) physical capability, whereby pregnant women were fit to contract the affected muscles to perform PFMT, and (2) psychological capability signifying pregnant women understood the correct method of performing PFMT.

Opportunity also had two categories: (1) physical opportunity, such as the KEPT-app itself, and (2) social opportunity in which pregnant women were able to understand further and perform PFMT at their own time. Meanwhile, motivation has relationship with cognitive ability, which boosted women's confidence to perform PFMT. There were two types of motivation: (1) reflective motivation in which pregnant women incorporated their thought processes to arrange PFMT to be done three times daily, and (2) automatic motivation in which the pregnant women adopted PFMT as part of their routine.

Additionally, the COM-B was integrated with the persuasiveness of the app (Table 5). The KEPT app should be tailored (primary task support) based on the intensity of the exercise, and users have the opportunity to self-monitor (primary task support). The app should be able to send reminders (dialogue support) to remind the user to perform the exercise at a certain time. The expertise and authority (system-credibility support) involved in developing the app are available in the video to convince the app user. Finally, an app should be designed and developed from credible and trustworthy sources to bolster users' confidence and trust.

Table 5. KEPT[a] app COM-B[b] model with persuasive system design.

COM-B model and features of the mHealth[c] app	Persuasive system design
Capability	
Psychological	
Educational video by a registered physiotherapist with an example patient	System credibility-expertise and authority
Physical	
Training timer according to the user's confidence and capability.	Primary support-tailoring.
Opportunity	
Physical	
The KEPT app was produced by our local University	System credibility-trustworthiness
Social	
Frequent Asked Question (FAQs) to provide further information	Primary support- Tailoring
Motivation	
Reflective	
Improve the understanding of the risks of pelvic floor muscle weakness by watching the video.	System credibility-expertise and authority
Calendar charting of the UI[d] symptoms	Primary task-self-monitoring
Automatic	
Daily reminder to perform PFMT[e] as their routine behavior.	Dialogue support-reminder

[a]KEPT: Kegel Exercise Pregnancy Training.

[b]COM-B: capability, opportunity, motivation behavior-model.

[c]mHealth: mobile health.

[d]UI: urinary incontinence.

[e]PFMT: pelvic floor muscle training.

Formative Research (Focus Group Discussion)

FGDs were conducted among 24 pregnant women at two primary care clinics to understand the desirability of the app design. The participants were invited via purposive sampling while waiting for their modified oral glucose tolerance test. A total of 5 sessions were conducted as listed in Table 6.

These interviews were conducted from September to November 2019. The discussion was conducted in a deductive manner to understand the most preferred user interfaces. The questions aimed to make the app as simple as possible since the participants will be using the app 3 times every day. The low-fidelity app designs were provided, and participants could select either Figure 3 (with 6 user interfaces) or Figure 4 (with 4 user interfaces).

Study participants were being asked about their experiences adhering to the exercise:

> *I have an experience with the PFMT. After I gave birth, I was selected to be followed up by a physiotherapist. The physiotherapist instructed me to perform the PFMT. He then inserted a camera and showed me the muscles contracted. I just performed the exercise. I just know how to perform it, and I do as the physiotherapist instructed to achieve 100 times a day. I did not count it as I do it regularly everyday all the time.* [Participant #18]

The statement suggested that pregnant women with or without urinary symptoms were motivated to perform and adhere to the exercise even after delivery. A correct understanding of PFMT importance assisted in compliance with daily exercise.

Following this, they were asked to share their opinion regarding adding PFMT notes into the KEPT app. The majority of study participants (19/24, 79.1%) preferred the apps without notes due to the time factor. However, one participant disagreed and mentioned:

> *…but it is better to have both notes and video. Sometimes, I want to know more about the exercise…* [Participant #20]

The response suggested mixed opinions on whether or not to include additional notes on the exercise.

Finally, regarding the design selection of the app, all the study participants from the FGD preferred design 2 (a more straightforward and minimalist concept). A high-fidelity prototype design was developed based on all the above findings as illustrated in Figure 5.

Table 6. Study participants' characteristics and opinions.

ID	Age	Ethnicity	Occupation	UI[a]	PFMT[b]	Timer	FAQ[c]
01	29	Malay	Ex-document manager	Yes	Yes	Yes	No
02	25	Malay	Pharmacist	Yes	No	Yes	No
03	24	Malay	Housewife	No	No	Yes	No
04	21	Malay	Housewife	No	No	Yes	No
05	25	Malay	Graph designer	No	No	Yes	No
06	29	Malay	Housewife	No	No	Yes	No
07	31	Chinese	Engineer	No	No	Yes	No
08	31	Malay	Clerk	No	No	Yes	No
09	31	Malay	Clerk	No	No	Yes	No
10	28	Malay	Clerk	No	No	Yes	No
11	25	Malay	Housewife	No	No	Yes	No
12	23	Indian	Housewife	Yes	No	Yes	No
13	31	Malay	Admin	Yes	No	Yes	No
14	24	Malay	Ex-Assistant Pharmacist	Yes	No	Yes	No
15	27	Malay	Housewife	Yes	No	Yes	No
16	27	Malay	Admin	No	No	Yes	Yes
17	30	Malay	Housewife	No	No	Yes	Yes
18	33	Malay	Staff Nurse	No	Yes	Yes	Yes
19	27	Malay	Housewife	No	No	Yes	Yes
20	21	Malay	Housewife	No	No	Yes	Yes
21	24	Malay	Make-up artist	No	No	Yes	No
22	22	Malay	Housewife	No	No	Yes	No
23	19	Malay	Salesperson	No	No	Yes	No
24	25	Malay	Housewife	No	No	Yes	No

[a]UI: urinary incontinence.

[b]PFMT: pelvic floor muscle exercise.

[c]FAQ: frequently asked questions.

Figure 3. Low-fidelity design of the KEPT app (six user interfaces). KEPT: Kegel Exercise Pregnancy Training.

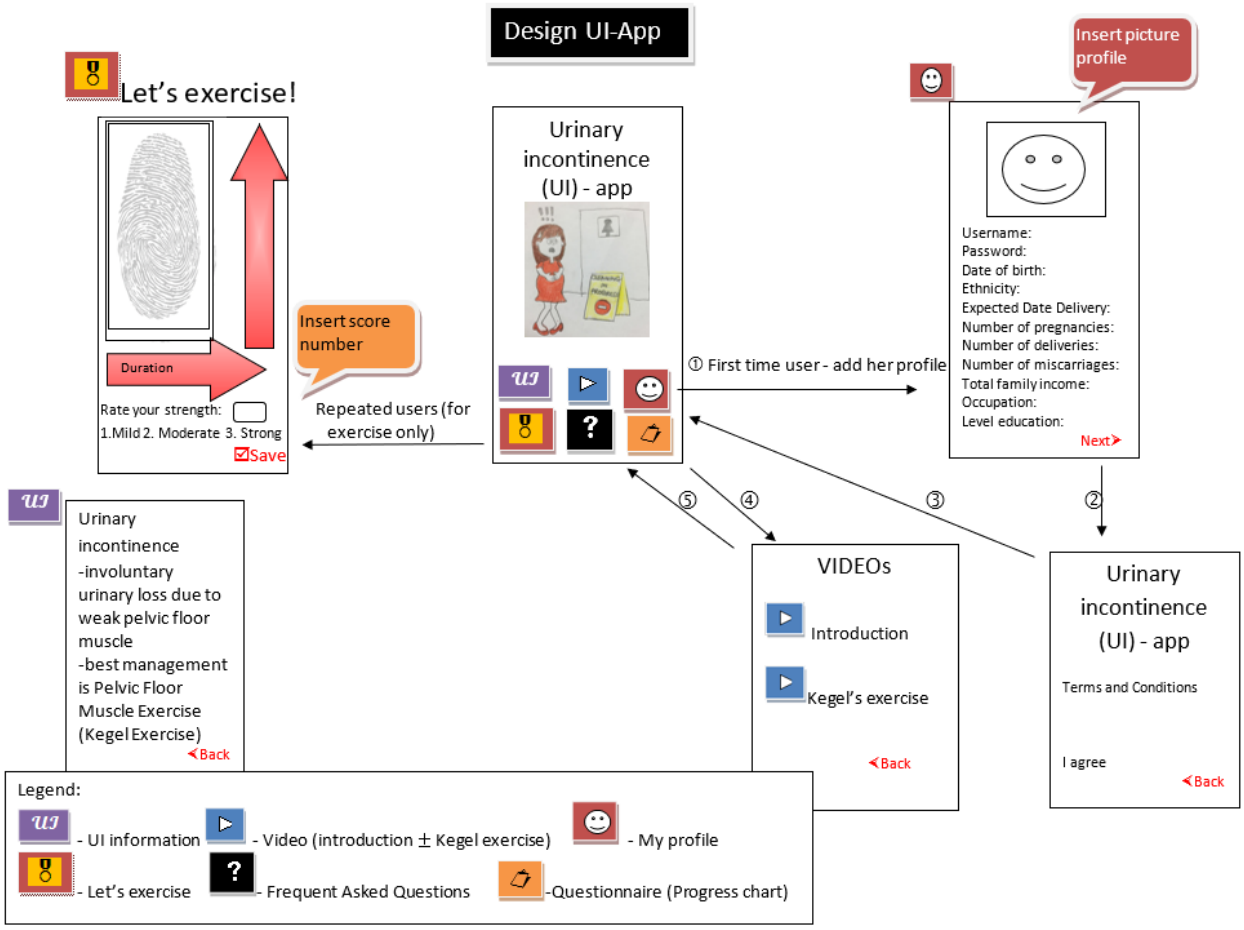


Figure 4. Low-fidelity design of the KEPT app (four user interfaces). KEPT: Kegel Exercise Pregnancy Training.

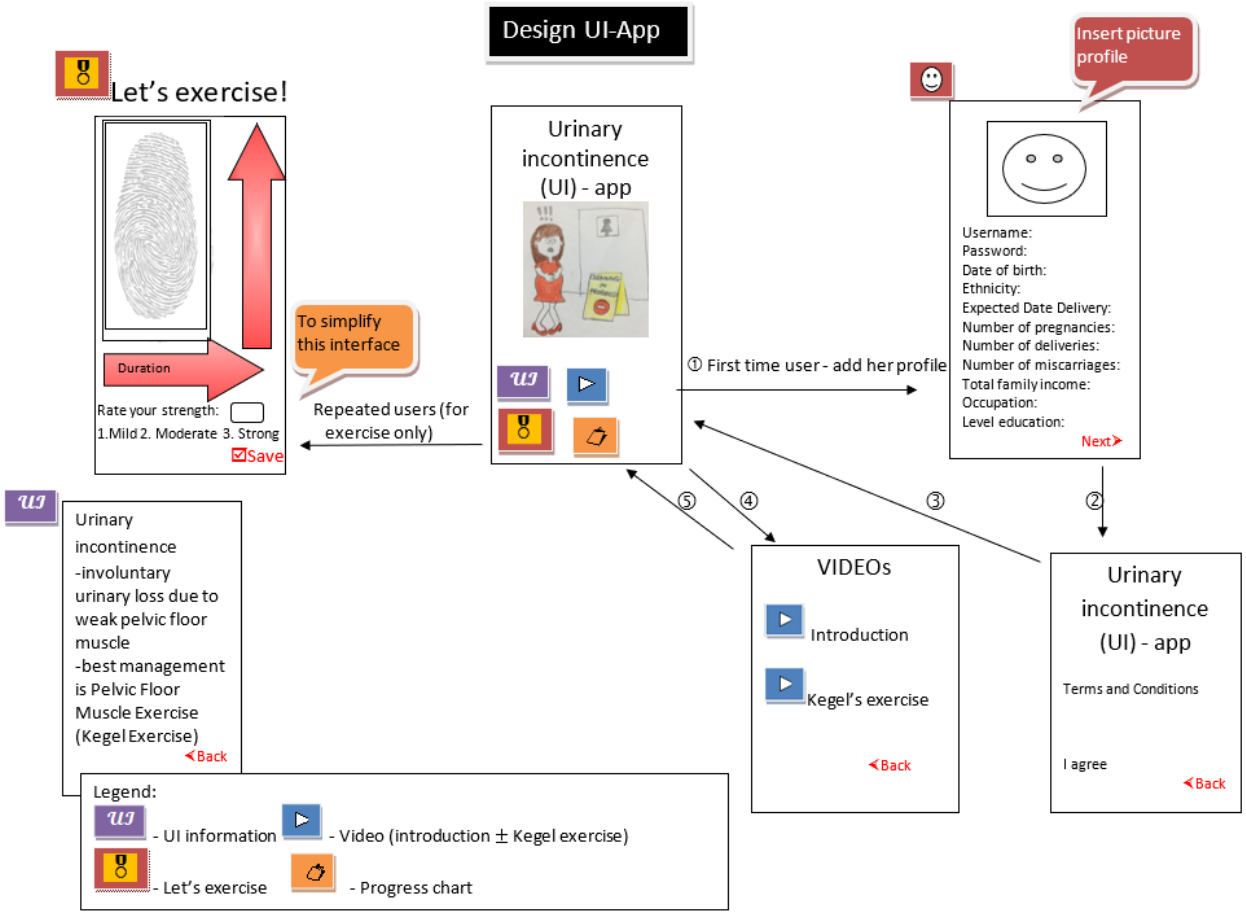
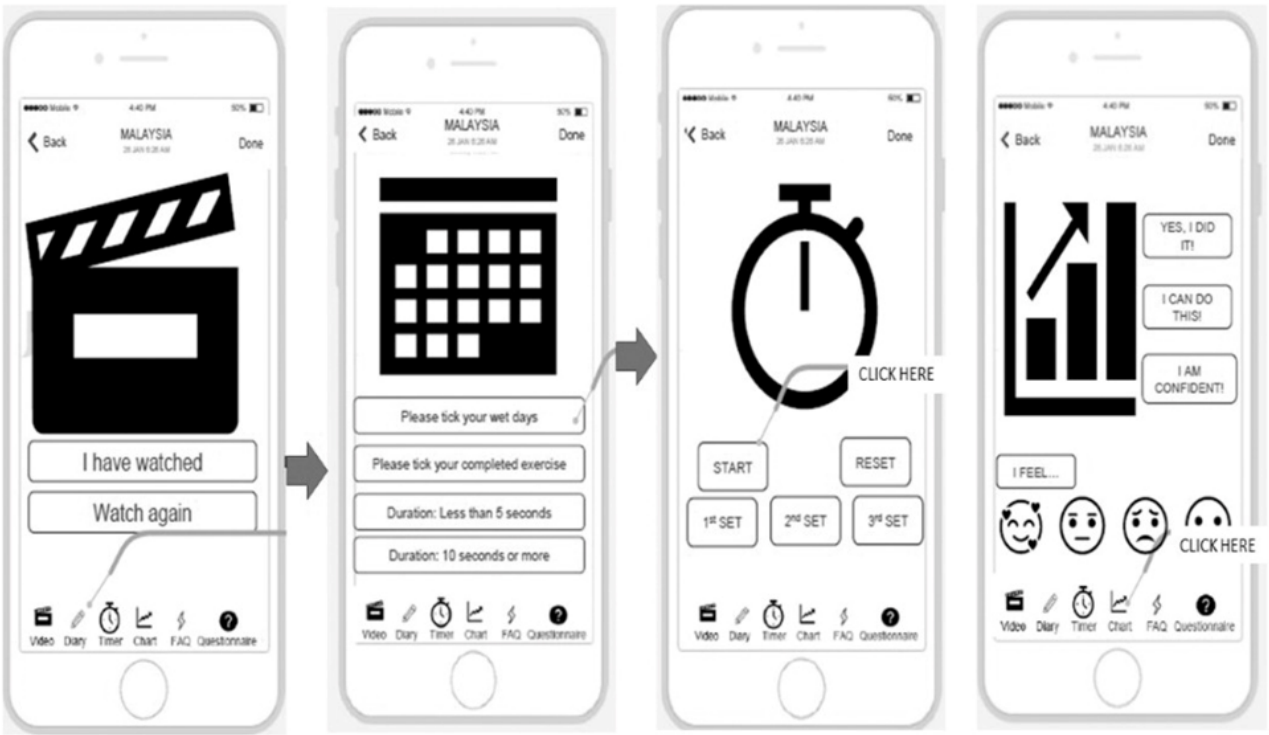


Figure 5. High-fidelity prototype design. KEPT: Kegel Exercise Pregnancy Training.



Prototype Development

The KEPT app prototype was proposed and discussed with the developer. The developer has designed the prototype with a blue color background [37] as illustrated in Figure 6. It consisted of both languages: Malay and English, whereby common English words were used. A minimum number of words were used to ensure the minimalist concept as target users were busy taking care of their families and/or working. There were 2 timer sounds: a high-pitch sound for exercise and a low-pitch sound during 6 seconds of rest. The reminder will be delivered if the PFMT is not completed at 6 pm on the same day.

The KEPT app interface did not require participants to log out due to the need to train three times daily. The UCD-11 checklist has been used as a guide for developing the KEPT app according to its user-centeredness properties as listed in Table 7.

Figure 6. Prototype KEPT app version 1.0. KEPT: Kegel Exercise Pregnancy Training.

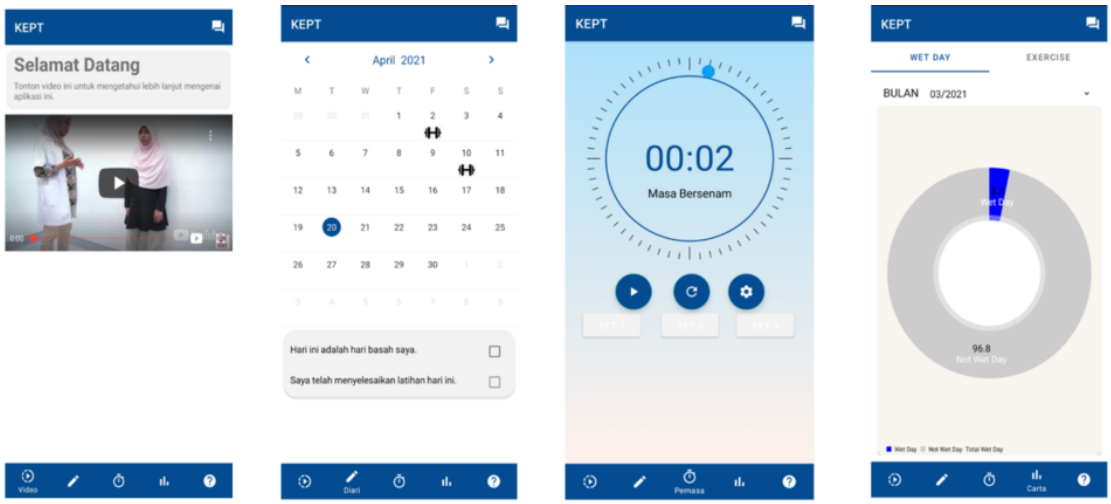


Table 7. UCD-11[a] Items in the KEPT[b] app

UCD-11 item	KEPT app
Were potential end users (eg, patients, caregivers, family and friends, and surrogates) involved in any steps to help understand users (eg, who they are, in what context might they use the tool) and their needs?	Pregnant women with urinary incontinence are involved in the needs assessment. They will use the app as a supportive tool for self-empowerment to improve their pelvic floor muscle strength.
Were potential end users involved in any steps of designing, developing, and/or refining a prototype	A cross-sectional study to understand the needs assessment [43].
Were potential end users involved in any steps intended to evaluate prototypes or a final version of the tool?	Study protocol for a randomized control trial has been published for this evaluation [51].
Were potential end users asking their opinions of the tool in any way?	Users' usability testing study has been completed and is currently under the manuscript writing process.
Were potential end users observed using the tool in any way?	Users' usability evaluation (think aloud method) study has been completed and is currently under the manuscript writing process.
Did the development process have 3 or more iterative cycles?	1. First iterative was the focus group discussion. 2. Second iterative cycle was with the user's usability study. 3. The third iterative cycle is currently being conducted in a pilot feasibility study [52].
Were changes between iterative cycles explicitly reported in any way?	The users' usability study is undergoing its manuscript writing process.
Were health professionals asked their opinion of the tool at any point?	The researcher team includes a family medicine specialist, public health specialist, physiotherapist, and community health specialist involved during the development of the app.
Were health professionals consulted before the first prototype was developed?	A family medicine specialist, public health specialist, and physiotherapist were consulted before the first prototype was developed.
Were health professionals consulted between initial and final prototypes?	The research team includes a family medicine specialist, public health specialist, physiotherapist, and community health specialist
Was an expert panel involved?	The research team and the software developer were involved in the development of the app.

[a]UCD-11: user-centeredness design-11.

[b]KEPT: Kegel Exercise Pregnancy Training.

Discussion

The KEPT app was designed and developed from an IM approach integrated with the mHealth development and evaluation framework. The COM-B model combined with PSD may help to improve the target audience's "trust" towards the app. Prior to this study, there were few PFMT apps available on the iOS and Android platforms. However, most of the apps had minimum specific strategies for enhancing adherence [53].

The cross-sectional study showed a significant association between UI and knowledge of PFMT (P=.01) and attitude towards PFMT (P=.01), with poor PFMT practice despite participants' knowledge was good [43]. This signified that not only are the training techniques but knowledge and attitude are crucial in managing UI among pregnant women. This result was supported by the FGDs, where one group highly insisted on having brief notes regarding PFMT via illustration (ie, the anatomy, physiology, and correct techniques).

Additionally, brief notes or a version of FAQs will serve as an information accuracy checklist, which is included in one of the app trustworthiness checklists [54]. Trustworthiness implies the qualities an individual requires to consider the app as trustable and consists of information accuracy, transparency, organizational attributes, societal influences, and external pressure [54]. Being able to deliver trustable and accurate PFMT techniques will improve PFMT self-efficacy among pregnant women, resulting in improved adherence to PFMT.

Pregnant women in this study also stated that they were unaware of the importance of incorporating good design into mHealth apps monitoring their daily activities, for example, recording the baby's movements. Nevertheless, with all the actionable features, the KEPT app was designed to enable pregnant women to adopt PFMT as their new instigation habit. Perhaps, in the future, the KEPT app may be upgraded by adding an additional antenatal diary log that may consist of a fetal movement chart, dietary nutrition, and physical activities.

COM-B was used to select the correct intervention (ie, training the pelvic floor muscle) based on previous studies [43,44]. The PSD component of the system's credibility and trustworthiness [55], with the expertise involved in the development, may add to the user's sense of safety and reliability regarding the KEPT app. Additionally, the KEPT app's reminder, self-monitoring, and PFMT timer were discussed. These three components may assist pregnant women in signaling the environment, whereby it is a habit trigger to get prepared for PFMT and go into an automatic mode [56]. Once the habit has been established, a person will be inclined to perform the behavior unconsciously or effortlessly with minimum awareness [57].

The KEPT app was the first app developed for a pregnancy-related target audience from the UCD approach, improving its acceptability and engagement [52]. There was an implication of applying the UCD-11 checklist as it is systematic and comprehensive, which will assist future researchers in developing the mHealth app effectively. However, being iterative for at least three times may have added challenges and financial complications to comply with. The prototype repeated needs in terms of to be evaluation, redevelopment, and re-evaluation may demotivate the researchers and software developers to undergo the iteration again. The KEPT app is currently undergoing pilot testing before entering the randomized control trial phase [51].

This study has a few limitations, such as time constraints and movement restriction orders. Although the study was conducted for an appropriate duration, curfew and restriction movement orders impacted the documentation, administration, and development of the software due to the COVID-19 pandemic situation in Malaysia.

Conclusions

The KEPT app was developed from a UCD-based behavioral change theory and accompanied by the PSD to improve users' engagement. The integration of the UCD-11 checklist with COM-B and PSD has prevailed to benefit the target user effectively. Therefore, it is crucial that the targeted users evaluate the usability and user acceptance of the final prototype in our next study.

Acknowledgments

The authors would like to thank the Director-General of Health Malaysia for permission to publish this paper. We would like to acknowledge the clinics' staff nurses, sisters, and matron involved in assisting the FGDs. The authors would like to thank the respondents involved in the study. This research was funded by the Geran Putra Berimpak Universiti Putra Malaysia (grant number UPM/800—3/3/1/GPB/2018/9668500).

Authors' Contributions

AJ contributed to the conceptualization, methodology, visualization, and formal analysis of the study, including resources and data curation. AJ and NDI contributed to the investigation. AJ, NA, and SNAS drafted and revised the original manuscript. SMS, NA, SNAS, and CNF revised and edited the manuscript. SMS supervised the study, and AJ and CNF administered the project. SMS acquired the necessary funding. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Abrams P, Andersson K, Apostolidis A, Birder L, Bliss D, Brubaker L, members of the committees. 6th International Consultation on Incontinence. Recommendations of the International Scientific Committee: Evaluation and Treatment of Urinary Incontinence, Pelvic Organ Prolapse and Faecal Incontinence. Neurourol Urodyn 2018 Sep 14;37(7):2271-2272. [doi: 10.1002/nau.23551] [Medline: 30106223]
2. Mostafaei H, Sadeghi-Bazargani H, Hajebrahimi S, Salehi-Pourmehr H, Ghojazadeh M, Onur R, et al. Prevalence of female urinary incontinence in the developing world: A systematic review and meta-analysis-A Report from the Developing World Committee of the International Continence Society and Iranian Research Center for Evidence Based Medicine. Neurourol Urodyn 2020 Apr;39(4):1063-1086. [doi: 10.1002/nau.24342] [Medline: 32243626]
3. Sangsawang B. Risk factors for the development of stress urinary incontinence during pregnancy in primigravidae: a review of the literature. European Journal of Obstetrics & Gynecology and Reproductive Biology 2014 Jul;178:27-34. [doi: 10.1016/j.ejogrb.2014.04.010]
4. Falah-Hassani K, Reeves J, Shiri R, Hickling D, McLean L. The pathophysiology of stress urinary incontinence: a systematic review and meta-analysis. Int Urogynecol J 2021 Mar;32(3):501-552 [FREE Full text] [doi: 10.1007/s00192-020-04622-9] [Medline: 33416968]
5. Moossdorff-Steinhauser HFA, Berghmans BCM, Spaanderman MEA, Bols EMJ. Prevalence, incidence and bothersomeness of urinary incontinence in pregnancy: a systematic review and meta-analysis. Int Urogynecol J 2021 Jul;32(7):1633-1652 [FREE Full text] [doi: 10.1007/s00192-020-04636-3] [Medline: 33439277]
6. Maeda N, Urabe Y, Suzuki Y, Hirado D, Morikawa M, Komiya M, et al. Cross-Sectional Study of the Prevalence and Symptoms of Urinary Incontinence among Japanese Older Adults: Associations with Physical Activity, Health-Related Quality of Life, and Well-Being. IJERPH 2021 Jan 06;18(2):360. [doi: 10.3390/ijerph18020360]
7. Przydacz M, Skalski M, Sobanski J, Chlosta M, Raczynski K, Klasa K, et al. Association between Lower Urinary Tract Symptoms and Sleep Quality of Patients with Depression. Medicina 2021 Apr 19;57(4):394. [doi: 10.3390/medicina57040394]
8. Al Kiyumi MH, Al Belushi ZI, Jaju S, Al Mahrezi AM. Urinary Incontinence Among Omani Women: Prevalence, risk factors and impact on quality of life. Sultan Qaboos Univ Med J 2020 Feb;20(1):e45-e53 [FREE Full text] [doi: 10.18295/squmj.2020.20.01.007] [Medline: 32190369]
9. Wang K, Xu X, Jia G, Jiang H. Risk Factors for Postpartum Stress Urinary Incontinence: a Systematic Review and Meta-analysis. Reprod Sci 2020 Dec;27(12):2129-2145. [doi: 10.1007/s43032-020-00254-y] [Medline: 32638282]
10. Woodley SJ, Lawrenson P, Boyle R, Cody JD, Mørkved S, Kernohan A, et al. Pelvic floor muscle training for preventing and treating urinary and faecal incontinence in antenatal and postnatal women. Cochrane Database Syst Rev 2020 May 06;5:CD007471. [doi: 10.1002/14651858.CD007471.pub4] [Medline: 32378735]
11. Sobhgol SS, Smith CA, Dahlen HG. The effect of antenatal pelvic floor muscle exercises on labour and birth outcomes: a systematic review and meta-analysis. Int Urogynecol J 2020 Nov;31(11):2189-2203. [doi: 10.1007/s00192-020-04298-1] [Medline: 32506232]
12. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
13. Salmon VE, Hay-Smith EJC, Jarvie R, Dean S, Terry R, Frawley H, et al. Implementing pelvic floor muscle training in women's childbearing years: A critical interpretive synthesis of individual, professional, and service issues. Neurourol Urodyn 2020 Feb;39(2):863-870 [FREE Full text] [doi: 10.1002/nau.24256] [Medline: 31845393]
14. Grant A, Currie S. Qualitative exploration of the acceptability of a postnatal pelvic floor muscle training intervention to prevent urinary incontinence. BMC Womens Health 2020 Jan 17;20(1):9 [FREE Full text] [doi: 10.1186/s12905-019-0878-z] [Medline: 31952500]
15. Moossdorff-Steinhauser HFA, Berghmans BCM, Spaanderman MEA, Bols EMJ. Urinary incontinence during pregnancy: prevalence, experience of bother, beliefs, and help-seeking behavior. Int Urogynecol J 2021 Mar;32(3):695-701 [FREE Full text] [doi: 10.1007/s00192-020-04566-0] [Medline: 33078344]
16. Bayat M, Eshraghi N, Naeiji Z, Fathi M. Evaluation of Awareness, Adherence, and Barriers of Pelvic Floor Muscle Training in Pregnant Women: A Cross-sectional Study. Female Pelvic Med Reconstr Surg 2021;27(1):e122-e126. [doi: 10.1097/spv.0000000000000852]
17. Dumoulin C, Hay-Smith J, Frawley H, McClurg D, Alewijnse D, Bo K, International Continence Society. 2014 consensus statement on improving pelvic floor muscle training adherence: International Continence Society 2011 State-of-the-Science Seminar. Neurourol Urodyn 2015 Sep;34(7):600-605. [doi: 10.1002/nau.22796] [Medline: 25998603]
18. Kay M, Santos J, Takane M. mHealth: New horizons for health through mobile technologies. mHealth: New horizons for health through mobile technologies. World Heal Organ. Geneva, Switzerland: ;64(7)?71; 2011. URL: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjZyPyn9Mz1AhWDT2wGHe4oBrMQFnoECAYQAQ&url=https%3A%2F%2Fwww.who.int%2Fgoe%2Fpublications%2Fgoe_mhealth_web.pdf&usg=AOv Vaw0Ugz0yLxYX4J7uCXzyDV7i [accessed 2021-11-30]
19. National Institute for Health Care Excellence (NICE). Evidence Standards Framework for Digital Health Technologies. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/corporate/ecd7 [accessed 2021-11-30]

20. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach. J Med Internet Res 2017 Apr 19;19(4):e125. [doi: 10.2196/jmir.5604]
21. Baharuddin R, Singh D, Razali R. Usability Dimensions for Mobile Applications-A Review. RJASET 2013 Feb 21;11(9):2225-2231. [doi: 10.19026/rjaset.5.4776]
22. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The Usability and Effectiveness of Mobile Health Technology-Based Lifestyle and Medical Intervention Apps Supporting Health Care During Pregnancy: Systematic Review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
23. Hauser-Ulrich S, Künzli H, Meier-Peterhans D, Kowatsch T. A Smartphone-Based Health Care Chatbot to Promote Self-Management of Chronic Pain (SELMA): Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth 2020 Apr 3;8(4):e15806. [doi: 10.2196/15806]
24. Vickery M, van Teijlingen E, Hundley V, Smith G, Way S, Westwood G. Midwives' views towards women using mHealth and eHealth to self-monitor their pregnancy: A systematic review of the literature. Eur J Midwifery 2020 Sep 17;4(September):1-11. [doi: 10.18332/ejm/126625]
25. Soltani H, Furness P, Arden M, McSeveny K, Garland C, Sustar H, et al. Women's and Midwives' Perspectives on the Design of a Text Messaging Support for Maternal Obesity Services: An Exploratory Study. J Obes 2012;2012:835464 [FREE Full text] [doi: 10.1155/2012/835464] [Medline: 22900153]
26. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
27. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
28. Osborne CL, Juengst SB, Smith EE. Identifying user-centered content, design, and features for mobile health apps to support long-term assessment, behavioral intervention, and transitions of care in neurological rehabilitation: An exploratory study. British Journal of Occupational Therapy 2020 Oct 07;84(2):101-110. [doi: 10.1177/0308022620954115]
29. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, Implementing, and Evaluating Mobile Health Technologies for Managing Chronic Conditions in Older Adults: A Scoping Review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
30. Cornet V, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold Stories in User-Centered Design of Mobile Health: Practical Challenges and Strategies Learned From the Design and Evaluation of an App for Older Adults With Heart Failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
31. Akmal Muhamat N, Hasan R, Saddki N, Mohd Arshad MR, Ahmad M. Development and usability testing of mobile application on diet and oral health. PLoS One 2021;16(9):e0257035 [FREE Full text] [doi: 10.1371/journal.pone.0257035] [Medline: 34495979]
32. Witteman HO, Vaisson G, Provencher T, Chipenda Dansokho S, Colquhoun H, Dugas M, et al. An 11-Item Measure of User- and Human-Centered Design for Personal Health Tools (UCD-11): Development and Validation. J Med Internet Res 2021 Mar 16;23(3):e15032 [FREE Full text] [doi: 10.2196/15032] [Medline: 33724194]
33. Sen S, Patel M, Sharma A. Software Development Life Cycle Performance Analysis. In: Mathur R, Gupta CP, Katewa V, Jat DS, Yadav N. editors. Singapore: Springer Singapore; Sep 28, 2021:311.
34. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
35. Oinas-Kukkonen H, Harjumaa M. A Systematic Framework for DesigningEvaluating Persuasive Systems BT - Persuasive Technology. In: Oinas-Kukkonen H, Hasle P, Harjumaa M, Segerståhl K, Øhrstrøm P, editors. Lecture Notes in Computer Science, vol 5033. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008:164-176. [doi: 10.1007/978-3-540-68504-3_15]
36. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive Technology in Mobile Applications Promoting Physical Activity: a Systematic Review. J Med Syst 2016 Mar;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
37. Jaffar A, Sidik SM, Admodisastro N, Mansor EI, Fong LC. Expert's Usability Evaluation of the Pelvic Floor Muscle Training mHealth App for Pregnant Women. IJACSA 2021;12(10):165-173. [doi: 10.14569/ijacsa.2021.0121019]
38. Kok G, Schaalma H, Ruiter RAC, van Empelen P, Brug J. Intervention mapping: protocol for applying health psychology theory to prevention programmes. J Health Psychol 2004 Jan;9(1):85-98. [doi: 10.1177/1359105304038379] [Medline: 14683571]
39. Fernandez ME, Ruiter RAC, Markham CM, Kok G. Intervention Mapping: Theory- and Evidence-Based Health Promotion Program Planning: Perspective and Examples. Front Public Heal. (AUG) 2019;7.
40. Green L, Marshall W. Kreuter: Health Program Planning: An educational and ecological approach. New York, USA: MC Graw Hill companies; 2005. p. A URL: http://www.lgreen.net/precede.htm
41. Sacomori C, Cardoso FL, Porto IP, Negri NB. The development and psychometric evaluation of a self-efficacy scale for practicing pelvic floor exercises. Brazilian Journal of Physical Therapy; 2013. [doi: 10.1590/S1413-35552012005000104]

42. Newman-Beinart NA, Norton S, Dowling D, Gavriloff D, Vari C, Weinman JA, et al. The development and initial psychometric evaluation of a measure assessing adherence to prescribed exercise: the Exercise Adherence Rating Scale (EARS). Physiotherapy 2017;103(2):180-185. [doi: 10.1016/j.physio.2016.11.001]
43. Jaffar A, Mohd-Sidik S, Nien FC, Fu GQ, Talib NH. Urinary incontinence and its association with pelvic floor muscle exercise among pregnant women attending a primary care clinic in Selangor, Malaysia. Rosier PFWM, editor. PLoS One. Public Library of Science; 2020 ;15(7):e0236140.
44. Jaffar A, Mohd-Sidik S, Abd Manaf R, Foo C, Gan Q, Saad H. Quality of life among pregnant women with urinary incontinence: A cross-sectional study in a Malaysian primary care clinic. PLoS One 2021;16(4):e0250714 [FREE Full text] [doi: 10.1371/journal.pone.0250714] [Medline: 33909678]
45. Alagirisamy P, Mohd Sidik S. Pelvic Floor Muscle Exercises During and After Pregnancy. Serdang: Universiti Putra Malaysia Press; 2020.
46. Bo K, Berghmans B, Morkved S, Van Kampen M. Evidence-Based Physical Therapy for the Pelvic Floor-E-Book: Bridging Science and Clinical Practice, 2nd ed. London, UK: Elsevier Health Sciences; 2014.
47. McClurg D, Frawley H, Hay-Smith J, Dean S, Chen SY, Chiarelli P. Scoping review of adherence promotion theories in pelvic floor muscle training - 2011 ics state-of-the-science seminar research paper i of iv. Neurourol Urodyn. John Wiley & Sons, Ltd 2015;34(7):614.
48. Woodley SJ, Hay-Smith EJC. Narrative review of pelvic floor muscle training for childbearing women-why, when, what, and how. Int Urogynecol J 2021 May 05:1-10. [doi: 10.1007/s00192-021-04804-z] [Medline: 33950309]
49. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
50. Garnett C, Crane D, West R, Brown J, Michie S. Identification of Behavior Change Techniques and Engagement Strategies to Design a Smartphone App to Reduce Alcohol Consumption Using a Formal Consensus Method. JMIR Mhealth Uhealth 2015;3(2):e73 [FREE Full text] [doi: 10.2196/mhealth.3895] [Medline: 26123578]
51. Sidik SM, Jaffar A, Foo CN, Muhammad NA, Abdul Manaf R, Ismail SIF, et al. KEPT-app trial: a pragmatic, single-blind, parallel, cluster-randomised effectiveness study of pelvic floor muscle training among incontinent pregnant women: study protocol. BMJ Open 2021 Jan 12;11(1):e039076 [FREE Full text] [doi: 10.1136/bmjopen-2020-039076] [Medline: 33436465]
52. Jaffar A, Mohd SS, Foo CN, Muhammad NA, Abdul MR, Fadhilah ISI. Protocol of a Single-Blind Two-Arm (Waitlist Control) Parallel-Group Randomised Controlled Pilot Feasibility Study for mHealth App among Incontinent Pregnant Women. Int J Environ Res Public Health. 2021 ;18(9):4792.
53. Latorre GFS, de Fraga R, Seleme MR, Mueller CV, Berghmans B. An ideal e-health system for pelvic floor muscle training adherence: Systematic review. Neurourol Urodyn 2019 Jan;38(1):63-80. [doi: 10.1002/nau.23835] [Medline: 30375056]
54. van Haasteren A, Gille F, Fadda M, Vayena E. Development of the mHealth App Trustworthiness checklist. DIGITAL HEALTH 2019 Nov 21;5:205520761988646. [doi: 10.1177/2055207619886463]
55. Asklund I, Nyström E, Sjöström M, Umefjord G, Stenlund H, Samuelsson E. Mobile app for treatment of stress urinary incontinence: A randomized controlled trial. Neurourol Urodyn 2017 Jun;36(5):1369-1376. [doi: 10.1002/nau.23116] [Medline: 27611958]
56. Chen W, Chan T, Wong L, Looi C, Liao C, Cheng H, et al. IDC theory: habit and the habit loop. RPTEL 2020 May 14;15(1):1-19. [doi: 10.1186/s41039-020-00127-7]
57. Hagger M. Habit and physical activity: Theoretical advances, practical implications, and agenda for future research. Psychology of Sport and Exercise 2019 May;42:118-129. [doi: 10.1016/j.psychsport.2018.12.007]

Abbreviations

COM-B: capability, opportunity, and motivation-behavior
FAQ: frequently asked question
HCP: health care providers
IM: intervention mapping
KEPT: Kegel Exercise Pregnancy Training
mHealth: mobile health
PFME: pelvic floor muscle exercise
PFMT: pelvic floor muscle training
PMF: pelvic floor muscle
PSD: persuasive system design
QoL: quality of life
SLDC: software development life cycle
UCD: user-centered design
UI: urinary incontinence

Edited by A Kushniruk; submitted 04.06.21; peer-reviewed by B Chaudhry, A Joseph; comments to author 17.10.21; revised version received 10.11.21; accepted 30.11.21; published 03.02.22.

Please cite as:
Jaffar A, Mohd-Sidik S, Foo CN, Admodisastro N, Abdul Salam SN, Ismail ND
Improving Pelvic Floor Muscle Training Adherence Among Pregnant Women: Validation Study
JMIR Hum Factors 2022;9(1):e30989
URL: https://humanfactors.jmir.org/2022/1/e30989
doi:10.2196/30989
PMID:35113025

Original Paper

A Remote Patient-Monitoring System for Intensive Care Medicine: Mixed Methods Human-Centered Design and Usability Evaluation

Akira-Sebastian Poncette[1,2*], MD; Lina Katharina Mosch[1,2*], MD; Lars Stablo[3], MSc; Claudia Spies[2], MD, PhD; Monique Schieler[2]; Steffen Weber-Carstens[2], MD, PhD; Markus A Feufel[3], MSc, PhD; Felix Balzer[1], MSc, MD, PhD

[1]Institute of Medical Informatics, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[2]Department of Anesthesiology and Intensive Care Medicine, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[3]Division of Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Felix Balzer, MSc, MD, PhD
Institute of Medical Informatics
Charité – Universitätsmedizin Berlin
Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin
Charitéplatz 1
Berlin, 10117
Germany
Phone: 49 30450 ext 651166
Email: felix.balzer@charite.de

Abstract

Background: Continuous monitoring of vital signs is critical for ensuring patient safety in intensive care units (ICUs) and is becoming increasingly relevant in general wards. The effectiveness of health information technologies such as patient-monitoring systems is highly determined by usability, the lack of which can ultimately compromise patient safety. Usability problems can be identified and prevented by involving users (ie, clinicians).

Objective: In this study, we aim to apply a human-centered design approach to evaluate the usability of a remote patient-monitoring system user interface (UI) in the ICU context and conceptualize and evaluate design changes.

Methods: Following institutional review board approval (EA1/031/18), a formative evaluation of the monitoring UI was performed. Simulated use tests with think-aloud protocols were conducted with ICU staff (n=5), and the resulting qualitative data were analyzed using a deductive analytic approach. On the basis of the identified usability problems, we conceptualized informed design changes and applied them to develop an improved prototype of the monitoring UI. Comparing the UIs, we evaluated perceived usability using the System Usability Scale, performance efficiency with the normative path deviation, and effectiveness by measuring the task completion rate (n=5). Measures were tested for statistical significance using a 2-sample *t* test, Poisson regression with a generalized linear mixed-effects model, and the N-1 chi-square test. $P<.05$ were considered significant.

Results: We found 37 individual usability problems specific to monitoring UI, which could be assigned to six subcodes: usefulness of the system, response time, responsiveness, meaning of labels, function of UI elements, and navigation. Among user ideas and requirements for the UI were high usability, customizability, and the provision of audible alarm notifications. Changes in graphics and design were proposed to allow for better navigation, information retrieval, and spatial orientation. The UI was revised by creating a prototype with a more responsive design and changes regarding labeling and UI elements. Statistical analysis showed that perceived usability improved significantly (System Usability Scale design A: mean 68.5, SD 11.26, n=5; design B: mean 89, SD 4.87, n=5; $P=.003$), as did performance efficiency (normative path deviation design A: mean 8.8, SD 5.26, n=5; design B: mean 3.2, SD 3.03, n=5; $P=.001$), and effectiveness (design A: 18 trials, failed 7, 39% times, passed 11, 61% times; design B: 20 trials, failed 0 times, passed 20 times; $P=.002$).

Conclusions: Usability testing with think-aloud protocols led to a patient-monitoring UI with significantly improved usability, performance, and effectiveness. In the ICU work environment, difficult-to-use technology may result in detrimental outcomes for staff and patients. Technical devices should be designed to support efficient and effective work processes. Our results suggest that this can be achieved by applying basic human-centered design methods and principles.

Trial Registration: ClinicalTrials.gov NCT03514173; https://clinicaltrials.gov/ct2/show/NCT03514173

(JMIR Hum Factors 2022;9(1):e30655) doi:10.2196/30655

KEYWORDS

digital health; patient monitoring; intensive care medicine; intensive care unit; technological innovation; user-centered design; usability; user experience; implementation science; qualitative research; interview; mixed methods; mobile phone

Introduction

Background

Continuous monitoring of vital signs is essential for patient safety in the intensive care unit (ICU) and emergency room [1]. It is also becoming increasingly relevant in general wards [2]. In the past decade, particularly in the context of the digital transformation of health care, vital sign monitoring has undergone constant change and is being transformed and augmented by important technological innovations such as less invasive sensors, remote monitoring technology [3-5], and artificial intelligence for clinical decision support [6,7]. Together, these innovations hold great promise for improving patient safety and health care provision [8,9].

Effective implementation of novel technologies, such as remote patient-monitoring devices, faces a variety of barriers [10-12], including lack of adoption by clinicians, often because of poor usability of the respective technologies [13-15]. In addition to its importance in successful implementation, usability is closely related to the efficacy of the technology [16,17]. A lack of usability may lead to medical errors, thus compromising patient safety [18,19]. Therefore, usability evaluation and identification of specific usability problems are essential in the development of a novel technology and its implementation in the clinical setting. However, to date, usability problems remain prominent in health information technology (IT), suggesting that usability aspects are often neglected in the health IT development process [20-22].

The human-centered design (HCD) approach is centered on the involvement of end users and their experiences with the product throughout the design and development process [23]. Applying HCD in the early stages of the design of novel digital health technologies can improve usability, staff adoption, effectiveness, and efficiency [24,25]. Several frameworks and guidelines for redesigning health care interfaces in accordance with HCD have been published; however, their adoption in health care has been lagging, and evidence on the impact of this topic on clinical performance outcomes is scarce [26-32].

Aim

We aim to evaluate the usability of a remote patient-monitoring system and, specifically, identify usability problems, positive findings, and user ideas. We hypothesize that an HCD approach will help to implement evidence-based design changes that will improve the subjectively perceived usability and objective measures of the effectiveness and efficiency of the technology.

Methods

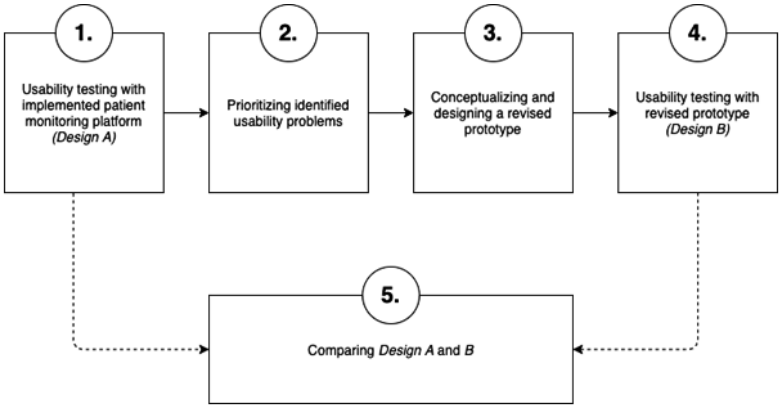
Ethics Approval and Consent to Participate

This study was approved by the ethics committee of the Charité–Universitätsmedizin Berlin (EA1/031/18). All participants provided consent before the study.

Study Design

Our usability study followed a five-step, mixed methods approach (Figure 1): (1) formative usability test of the implemented patient-monitoring platform interface design A [33], (2) identification and prioritization of usability problems, (3) conceptualization and design of prototype interface design B with informed design changes, (4) formative usability testing of design B, and (5) comparison of design A and design B. For usability testing, we applied simulated use tests with think-aloud protocols and performance measurements (subjectively perceived usability, efficiency, and effectiveness) [30,34]. For step 5, we chose a single-factor 2-group study design, as described by Gravetter and Forzano [35].

Figure 1. The research approach, beginning with usability testing and identification of major problems in design A, followed by prototyping of design B and its usability testing, concluding with a comparison between design A and design B.



Study Setting and Technical Setup

This study was conducted in the context of implementing the Vital Sync 2.4 virtual patient-monitoring platform (Medtronic plc) in the Post Anesthesia Care Unit, an ICU primarily for postoperative patients requiring short-term intensive care treatment and monitoring. VitalSync was used to monitor patients in the ICU from portable tablet computers on hospital premises. The primary patient-monitoring system used was the IntelliVue patient monitoring system (MX800 software, version M.00.03; MMS X2 software, version H.15.41-M.00.04) from Koninklijke Philips NV.

Between May 2018 and June 2019, the VitalSync monitoring system was installed for 5 of the 10 ICU beds. Two sensors (for pulse oximetry and capnography) recorded peripheral capillary oxygen saturation, pulse rate, end-tidal carbon dioxide, and respiratory rate at a frequency of 1 Hz. The VitalSync user interface (UI) was displayed on a monitor at the central station and on six tablet computers (2 standard iPads, 2 iPad minis, and 2 Microsoft Surfaces). The UI of the system was structured where the home screen gave an overview of patients admitted to the system, displayed in tiles (Figure 2). Displayed were numerical values for the monitoring parameters, the patient's name and bed location, and specific information on alarms if any. Clicking on a patient tile took the user to a screen with details about the selected patient (eg, graphical curves for end-tidal carbon dioxide values) and other functions (eg, displaying patient reports, linking, or unlinking devices). There was also the option of clicking on each parameter to see a trend analysis of that value. To link a patient to the system, the *Admit Patient* screen was accessed, and the patient ID was entered, after which the bed location and monitoring device could be selected to complete the admission process (Figure 3) [36-38]. Further technical description and details regarding the use of the software can be found elsewhere [10].

Figure 2. Home screen of the implemented patient-monitoring platform (design A). etCO2: end-tidal carbon dioxide; PR: pulse rate; RR: respiratory rate; SPO2: peripheral capillary oxygen saturation [36-38].

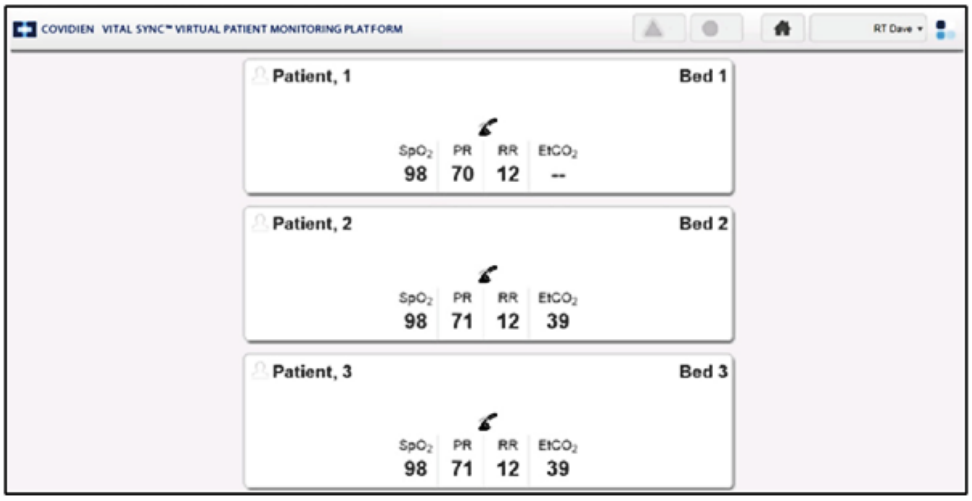
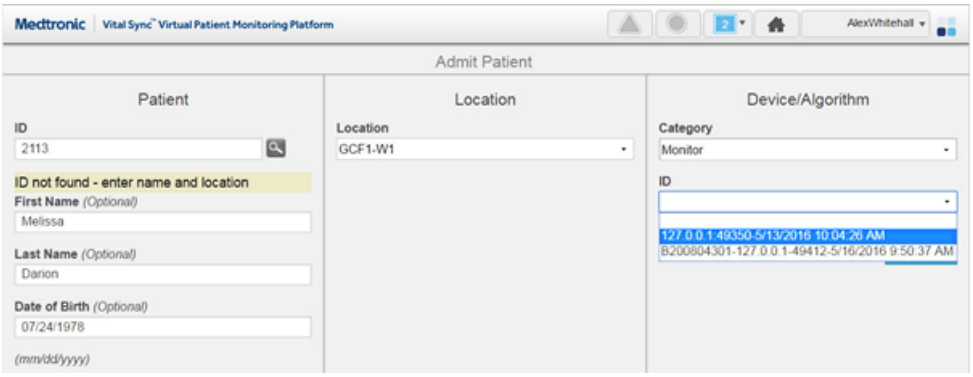


Figure 3. Admit Patient screen of the implemented patient-monitoring platform (design A) [36-38].



Research Team

Following the principles of HCD [39], our research team members have multidisciplinary skills and perspectives. Specifically, the team included a physician with a background in anesthesiology, intensive care medicine, geriatrics, and digital health (ASP); a senior medical student with a focus on digital health (LM); a senior human factors student with a background in engineering (LS); a professor of ergonomics with a PhD in human factors and industrial and organizational psychology (MF); the anesthesiology department's head of staff (CS); and a professor of medical data science, who is also a consultant anesthesiologist and computer scientist (FB).

Data Collection

Data collection took place from August 23, 2019, to March 10, 2020. Our data comprised think-aloud transcripts of the first block of usability tests (ie, design A), researcher notes (including

click patterns), and posttest questionnaires from the two blocks of usability tests (ie, design A and design B). We conducted 10 usability tests with ICU staff—5 (50%) tests each for design A (August and November 2019) and design B (February and March 2020). For recruitment, we contacted potential participants via email. We aimed to represent all professions working with the remote patient-monitoring system, namely anesthesiologists (3/10, 30%), ICU nurses (5/10, 50%), and respiratory therapists (2/10, 20%). Participation was voluntary, and no incentives were offered.

Usability testing of design A and design B was performed on an iPad mini 4 (model A1550). For testing sessions with design A, 5 patients in the ICU were connected to the system. This allowed real-time monitoring of the patients' vital signs on the iPad used by the participants. Testing of design B differed from testing of design A in that no patients were connected to actual sensors, and only one of the researchers was present during the testing sessions.

The testing sessions were conducted in German. Participants were asked about their profession and the number of years of professional experience in intensive care medicine. They were then given 4 tasks to complete while verbalizing their thoughts [40]. We provided the participants with the following use context: "A new patient was admitted to the unit and was connected to the $etCO_2$ and SpO_2 sensors (Mrs. Schmitt, born 01/01/1950, Patient-ID 12345, bed site 02)."

In accordance with the requirements for formative usability testing [41], participants were selected to complete the following key tasks during the simulated use test:

1. "Please add Mrs. Schmitt to the patients you want to monitor in Vital Sync™."
2. "You would like to see the trend of Mrs. Schmitt's oxygen saturation for the last two hours. How do you proceed?"
3. "You have identified that Mrs. Schmitt is actually not in bed 2 but in bed 6. You want to adjust this information in Vital Sync™. How do you proceed?"
4. "Mrs. Schmitt has been discharged. Please disconnect Mrs. Schmitt's devices and delete her entry from Vital Sync™."

Audio recordings of the simulated use tests were transcribed verbatim. A researcher who had not performed the transcription reviewed the transcripts. Immediately after the simulated use tests of both designs A and B, participants were asked to complete a posttest questionnaire, including the System Usability Scale (SUS) [42,43].

Data Analysis

Qualitative Analysis and Identification of Usability Problems

To analyze data from the think-aloud transcripts of design A testing sessions, we adapted a deductive analytic approach [44]. A coding scheme introduced by Kushniruk and Patel [44] was refined to the topic of study (patient monitoring in ICUs; Multimedia Appendix 1). Using the qualitative data analysis software MAXQDA 2018 (VERBI GmbH), think-aloud transcripts were coded according to the developed scheme. Coded segments (ie, usability problems) were specified into the subcodes, which were further summarized and listed (eg, meaning of labels unclear).

To decide which problems to eliminate first in the subsequent design iteration, summarized usability problems were ranked in terms of severity and frequency [45,46]. To assess problem severity, impact scores were assigned to each usability problem by 2 physicians who were experienced in intensive care medicine. The following scores were available for selection:

- The solution to this problem is subtle and possible enhancement or suggestion (score 1)
- The problem has a minor effect on usability (score 2)
- The problem creates significant delay and frustration (score 3)
- The problem prevents task completion (score 4)

Subsequently, the probability of occurrence was calculated by dividing the number of participants who encountered a particular problem by the total number of participants. To categorize problem frequency, each usability problem was assigned to one of four frequency levels: frequency ≤10% (level 1), frequency 11% to 50% (level 2), frequency 51% to 89% (level 3), and frequency ≥90% (level 4). Finally, criticality was calculated by adding the impact score and frequency levels [45] (eg, when a usability problem was rated as creating significant delays [impact score 3], which was experienced by 80% of participants [level 3], resulting in a criticality score of 6).

Analysis of Effectiveness, Efficiency, and Subjective Usability

The task completion rate [47,48] was measured to evaluate the effectiveness of design A and design B. Normative path deviation [49] was assessed based on participants' click patterns to account for efficiency. The sequence of steps users took when interacting with the interface to complete a task was compared with an optimal sequence of goal-directed steps defined by the researchers for each task. The difference between the normative path and observed path for each user and each task was calculated using the Levenshtein algorithm [33,49]. The SUS was used to assess the perceived usability of design A and design B [42,43,50].

Prototype Design

Design solutions were conceptualized by ASP and LS for all identified usability problems. This resulted in a list of ranked usability problems with the suggested design solutions. The identified usability problems from design A were revised by building design B, a clickable prototype, using Axure RP 9. A feedback loop was used to develop design B: one researcher (LS) built the prototype, and another researcher (ASP) reviewed the design and provided feedback from an intensivist's perspective.

Statistical Analysis

To assess the level of improvement between design A and design B, we hypothesized that the task completion rate for design B would be higher than that of design A, design B would lead to lower normative path deviation values than design A, and the SUS scores for design B would be higher than that of design A.

We used the N-1 chi-square test to compare the task completion rates of both designs [45]. To compare the normative path deviations for both designs, we used a Poisson regression drawing upon a generalized linear mixed-effects model with participants as random effects, as introduced by Schmettow et al [33]. A 2-sample *t* test was conducted to compare the SUS scores between design A and design B, as recommended by Sauro and Lewis [45]. We tested for normality using the Shapiro–Wilk test [51] and homoscedasticity (homogeneity of variance) using the Levene test [52].

Results

Overview and Sample

Measured by task completion rate, normative path deviation, and SUS score, design B was found to be significantly improved compared with design A. We first elaborate on the results of the qualitative analyses and then report the quantitative results.

The sample comprised a total of 10 ICU staff, aged 25 to 39 years, with work experience ranging from 1 to 20 years, who were divided into groups (5, 50% each) for the evaluation of the 2 designs.

Qualitative Results

Summary

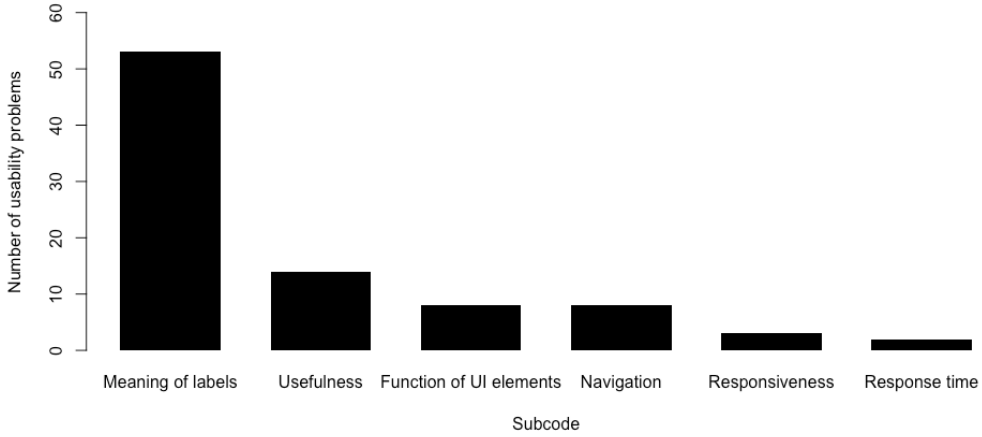
The coding of the transcripts revealed three main codes: usability problems, user ideas and requirements, and positive findings. The codes are visualized with a sunburst diagram (Figure 4; see Multimedia Appendix 1 for the adapted coding scheme by Kushniruk and Patel [44]). Items from the transcripts of the think-aloud protocols were mapped to the subcodes derived by Kushniruk and Patel [44] for the main categories—usability problems and positive findings. For usability problems, the items were assigned to the subcodes of usefulness of the system, response time, responsiveness, meaning of labels, function of UI elements, and navigation; for positive findings, the items were assigned to usefulness, overall ease of use, function of UI elements, layout/screen organization, and color.

Figure 4. Results of qualitative analysis of the think-aloud transcript. Three main codes were identified (inner ring) and subcoded (middle ring). The outer ring represents further information derived from the concrete items that were assigned to the subcodes (ie, specific user ideas or positive findings). UI: user interface.

Usability Problems

In total, 37 specific usability problems were identified (Multimedia Appendix 2). The number of usability problems related to the respective codes is visualized in Figure 5; most issues were related to labeling (53/88, 60%). The meaning of labels was mostly unclear—that is, participants were not familiar with certain terms (eg, the meaning of exclamation marks, abbreviations such as those for pulse rate [PR] and integrated pulmonary index [IPI], or terms such as *polardiagramm*). Users were concerned about whether a certain function was useful for the requirements of their clinical work or when a given task could not be accomplished (eg, participants selected the wrong bed site tile and participants were not sure about the correct patient or device ID; 14/88, 16%). There were difficulties in using or understanding the function of UI elements such as buttons (eg, gray circle or telescope symbols; 8/88, 9%). Furthermore, participants seemed to have problems navigating the monitoring system (ie, finding the right click path to admit patients to the platform; 8/88, 9%). Users criticized the responsiveness of the system (ie, the system did not behave as expected; 3/88, 3%) and the response time (ie, they complained about the time it took the device to respond; 2/88, 2%).

Figure 5. Number of occurrences for each subcode of usability problems. Meaning of labels (n=53), usefulness (n=14), function of UI elements (n=8), navigation (n=8), responsiveness (n=3), response time (n=2). UI: user interface.



User Ideas and Requirements

Users emphasized that the system's ease of use was particularly important to ensure its usability in emergency situations. The tool should be customizable to add other relevant vital signs (eg, intracranial pressure) or to display additional patient information. Participants required audible alarm notifications and the ability to share information regarding relevant patient events with colleagues (eg, about critical patient conditions). Vector graphics were suggested to allow zooming in and out of the vital sign curves. Moreover, participants demanded the ability to see curves of different parameters in an overlapping representation to be able to make inferences from one vital parameter to another. To facilitate spatial orientation, it was suggested that the beds be displayed in the UI according to the physical ward floor plan. Other ideas included adding a drag-and-drop function to rearrange multiple beds at once in the UI and integrating a high-frequency recording function to capture critical events.

Positive Findings

Participants stated that the system's scope of functionality was limited compared with other monitoring solutions. However, the reduced complexity was considered helpful in hospital wards with high patient turnover or stressful environments to get a quick overview of the patient's health condition. The system's mobility and overall ease of use were perceived as positive. Participants seemed to be familiar with the following basic UI elements: the home button depicted by a house, the editing symbol depicted by a pen, and the alarm symbol depicted by a warning triangle. Simplicity in the design and use of color was also rated as positive.

Design Iteration

The 37 distinct usability problems were ranked in relation to severity and frequency of occurrence (Multimedia Appendix 2). Potential solutions were assigned to the problems and were realized in design B (Figures 6 and 7). In total, 5 design iterations were performed between ASP and LS.

The main improvements in the prototype version compared with the previous interface were as follows:

- More responsive design
- Unknown labels were replaced or removed
- Unknown UI elements were replaced or removed
- A dashboard that counted beds, patients, and monitoring systems was added
- A confirmation dialog before replacing bed numbers was added
- State-of-the-art dark theme design was adapted from material.io

Figure 6. Redesign of the user interface of the prototype (design B) patient admission screen.

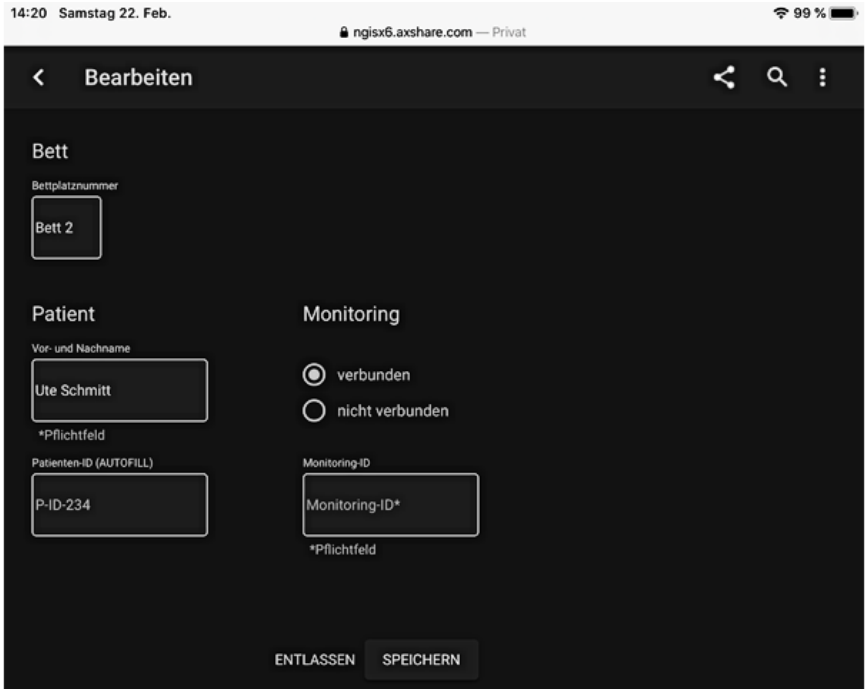
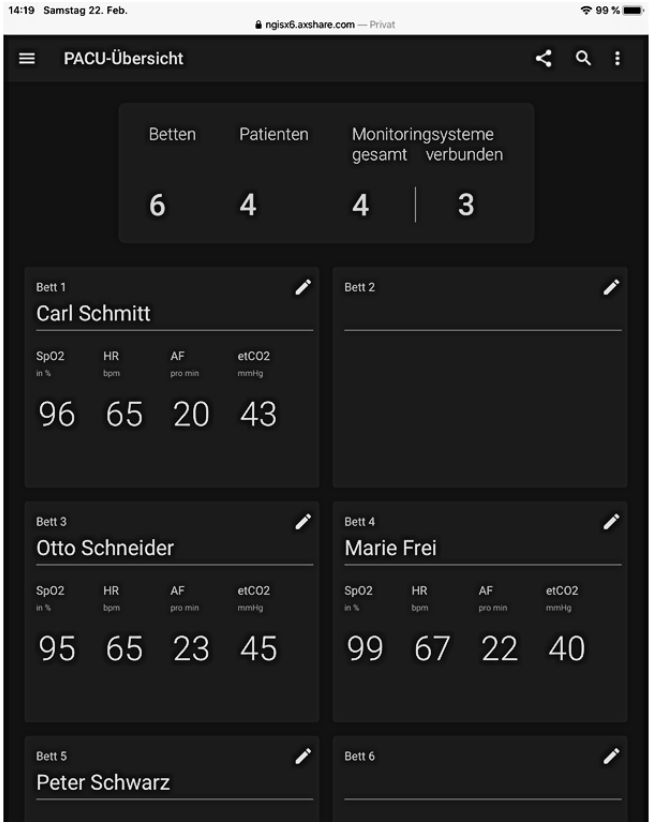


Figure 7. Redesign of the user interface of the prototype (design B) patient tile overview.



Quantitative Results

Effectiveness

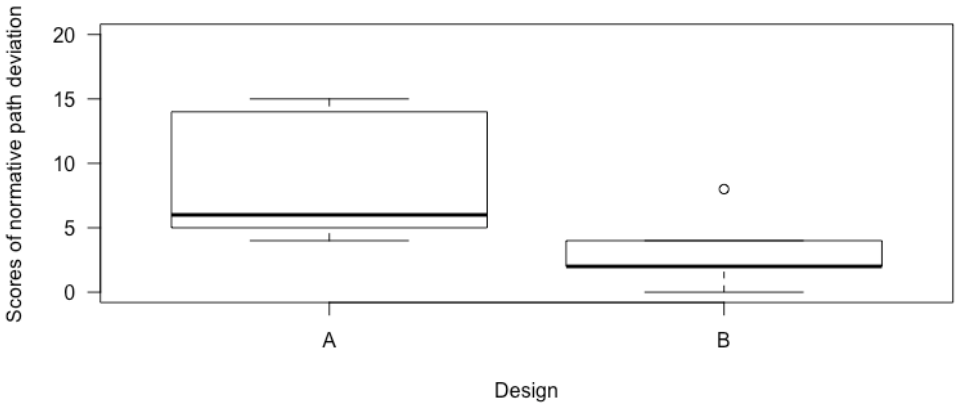
The task completion rate was higher for design B (attempts=20; 0/20, 0% failed and 20/20, 100% passed) than for design A (attempts=18; 7/18, 39% failed and 11/18, 61% passed). A 1-tailed N-1 chi-square test suggests that this is a statistically significant difference (χ^2_1=9.3; *P*=.002).

Efficiency

The average normative path deviation of design B (mean 3.2, SD 3.03; 5/10, 50%) was 63.4% lower than that of design A (mean 8.8, SD 5.26; 5/10, 50%; Figure 8). Poisson mixed-effects regression suggests that this reduction in the normative path

deviation is statistically significant ($\beta_{design\ B}$=−1.04, 95% CI −2.09 to −0.13; exp [$\beta_{design\ B}$]=1.13; *P*<.001).

Figure 8. Scores of normative path deviation for design A and design B. The circle symbolizes outliers. Outliers are defined in the box plots as values that have 1.5 times the distance between Q1 and Q3 (Q1 is the lower line of the box, and Q3 is the upper line of the box).



Usability

The average SUS score of design B (mean 89, SD 4.87; 5/10, 50%) was higher than that of design A (mean 68.50, SD 11.26; 5/10, 50%). This difference was statistically significant with a 1-tailed *t* test (t_8=3735; *P*=.003).

Discussion

Principal Findings

This study evaluated the usability of a remote patient-monitoring system (design A) by identifying the individual usability problems that informed the conceptualization and design of a revised prototype version (design B). Most of the usability problems identified were related to labeling, followed by the perceived lack of usefulness of the monitoring system [10,53,54]. The UI's navigation was frequently criticized by participants. Further identified usability problems include unclear UI elements, poor responsiveness, and increased response time. The resolution of the usability problems resulted in a significant increase in the perceived usability, efficiency, and effectiveness of the system.

Usability of Technologies in Intensive Care Medicine

Over the past 2 decades, the usability of health IT has been investigated in multiple studies applying different methodologies, revealing relatively poor usability and late involvement of end users in the development process [22,55]. This is reflected in our results; based on an HCD approach, we found a relatively high number of easy-to-solve usability problems, the resolution of which led to a significant improvement in the usability of the remote patient-monitoring solution. Most of the usability problems identified were related to labeling, an important issue that is addressed by regulatory requirements [30,56]. The UI's navigation was frequently criticized by participants. UI navigation problems can affect the overall usability of medical devices, especially in high-stress situations [57-59]. In this regard, simple, intuitive, and role-specific designs are beneficial [60-62], which is also reflected in the user ideas generated by the participants in our study.

The ICU is an exceptional environment that places diverse demands on health IT to be used there. High stress levels and patients who are unstable and critically ill, with varying care and treatment requirements, are among the conditions that must be considered [63-67]. Multiple digital devices already in place increase the cognitive load on staff as they are required to operate the devices and interpret their output [62,67]. Health care professionals applying physiological monitoring systems underuse the range of features currently available [28]. This might also be because of inadequate digital skills among health professionals and insufficient training of staff in the use of digital technologies [68-72].

With the increasing complexity and expanding the functionality of digital technologies and their increased use in all clinical settings, usability considerations have become all the more important to realize the full potential of such innovations. Given our findings, we suggest that HCD plays an important role in realizing the potential of IT in health care.

HCD in the Implementation of Digital Health Technology

Applying an HCD approach, the inclusion of usability testing and prototyping of a new UI for a remote patient-monitoring system increased usability, according to our findings. HCD encompasses the involvement of end users (ie, health care professionals) in the design and evaluation process, and the required efforts have been shown to be both worthwhile and beneficial in all development phases of a novel digital health technology, enhancing usability and performance [28,59,73]. Research suggests that user knowledge and beliefs about the technology to be implemented are key factors for the successful implementation of the technology [74]. Therefore, HCD should be applied not only during the design and development processes but also during implementation [55]. This could be achieved by establishing innovation and usability laboratories in universities and maximum care hospitals [75]. In the future, HCD is likely to be indispensable for improving both the performance and implementation of IT in health care.

Despite many publications demonstrating the benefits and relevance of usability testing and HCD in health care, there still seems to be a lack of awareness of its importance and the value

of involving key users in the early stages of technology development. The reasons for this may be the perceived costs and frequent lack of incentives to conduct usability evaluations. Moreover, as was the case in our study, the design and implementation of health technologies are often separate processes, making it difficult to apply an HCD approach across all development and implementation phases [22,73]. Further research needs to be conducted to explore how to overcome these barriers to obtain the most out of IT products in health care for both staff and patients.

Limitations

In this study, we showcase an HCD approach to improve the usability of a remote patient-monitoring system in a hospital setting. However, from a scientific perspective, there are several limitations to the scope of the study and the interpretation of results. Owing to the qualitative research design, it is not possible to quantify or generalize the usability problems identified to other health technologies and settings. In addition, translation of our results to other hospital settings or countries is limited because of the single-center design of this study and the relatively small sample size. It was not possible to draw samples randomly, which needs to be considered as a potential source of bias when interpreting the results. The comparison between design A, which was a working medical product installed in the ICU, and design B, a prototype mock-up, may be potentially unfair with a number of confounders in the 2 arms. Nonetheless, given the observed effects of meaningful labeling and easy-to-understand UIs on efficiency and effectiveness, our results help to underline the importance and potential of HCD for realizing the potential of IT in health care. Follow-up studies should be envisioned in collaboration with medical device manufacturers using design B.

We did not perform a usability test of all features of the remote patient-monitoring device, which comprises more than just the remote monitoring device UI (eg, sensors, bedside monitors, or cables are also part of it). We focused on tablet use for this study as it distinguishes remote patient monitoring from regular patient monitoring, and the tablet is the touchpoint with which the user interacts most frequently. Thus, we restricted the study scope to the UI of the tablet version of the remote monitoring system; that is, the smartphone and desktop UI versions were not investigated. We only tested the German version of the UI, which limits certain findings (eg, regarding the labeling) to German-speaking regions.

We were not able to refer to a standardized checklist or protocol for reporting the results of this study. The development of such a checklist or protocol could be an interesting area for further research, as it could improve the quality and reproducibility of usability study reports.

Conclusions

Applying an HCD approach with usability testing and conceptualized design of a revised prototype version significantly improved the usability of the remote patient-monitoring system for the end points of perceived ease of use, efficiency, and effectiveness. Technical devices should be designed to support efficient and effective work processes, especially in the sensitive working environment of the ICU, with usability being an essential facilitator of maximum performance, successful implementation, and ultimately patient safety. Our results suggest that HCD methods and principles can help realize the goals and potential of IT in health care. However, currently, HCD methods are often not applied early enough in the development process of digital health technologies for ICUs. Further research should explore how to increase early product evaluations in hospitals with end users to take better advantage of their input, not only for the development of user-friendly IT solutions but also for their successful implementation in clinical settings.

Acknowledgments

The authors express their gratitude to the intensive care unit staff for their participation in this study. The authors would like to thank the German Research Foundation (Deutsche Forschungsgemeinschaft), the Open Access publication fund of the Charité–Universitätsmedizin Berlin, and the Einstein Center Digital Future. ASP is a participant in the Charité Digital Clinician Scientist Program funded by the Charité–Universitätsmedizin Berlin and the Berlin Institute of Health. Finally, the authors would like to thank Medtronic for providing the aforementioned devices free of charge.

Authors' Contributions

CS had the idea for shared decision allocation and initiated the implementation of remote patient-monitoring in the intensive care unit. The study was conceived by ASP, CS, FB, and LKM. LS, ASP, and LKM conducted data acquisition and analysis. ASP and LKM wrote the manuscript, supported by LS. MAF supported the design of the study, conception of methodology, and interpretation of results. FB supervised all parts of the study. All authors critically reviewed and approved the manuscript. The paper was extracted from the MS thesis of LS.

Conflicts of Interest

CS and FB report funding from Medtronic. FB also reports grants from German Federal Ministry of Education and Research, grants from German Federal Ministry of Health, grants from Berlin Institute of Health, personal fees from Elsevier Publishing, grants from Hans Böckler Foundation, other from Robert Koch Institute, grants from Einstein Foundation, and grants from Berlin University Alliance outside the submitted work.

Multimedia Appendix 1
Coding scheme adapted from Kushniruk and Patel [44].
[PNG File , 55 KB - humanfactors_v9i1e30655_app1.png]

Multimedia Appendix 2
Usability ranking.
[XLSX File (Microsoft Excel File), 24 KB - humanfactors_v9i1e30655_app2.xlsx]

References

1. Moreno RP, Rhodes A, Donchin Y, European Society of Intensive Care. Patient safety in intensive care medicine: the Declaration of Vienna. Intensive Care Med 2009;35(10):1667-1672 [FREE Full text] [doi: 10.1007/s00134-009-1621-2] [Medline: 19697007]
2. Hu P, Galvagno Jr SM, Sen A, Dutton R, Jordan S, Floccare D, ONPOINT Group. Identification of dynamic prehospital changes with continuous vital signs acquisition. Air Med J 2014;33(1):27-33. [doi: 10.1016/j.amj.2013.09.003] [Medline: 24373474]
3. Balzer F, Habicher M, Sander M, Sterr J, Scholz S, Feldheiser A, et al. Comparison of the non-invasive Nexfin® monitor with conventional methods for the measurement of arterial blood pressure in moderate risk orthopaedic surgery patients. J Int Med Res 2016;44(4):832-843 [FREE Full text] [doi: 10.1177/0300060516635383] [Medline: 27142436]
4. Shah PS. Wireless monitoring in the ICU on the horizon. Nat Med 2020;26(3):316-317. [doi: 10.1038/s41591-020-0798-3] [Medline: 32161410]
5. Michard F, Bellomo R, Taenzer A. The rise of ward monitoring: opportunities and challenges for critical care specialists. Intensive Care Med 2019;45(5):671-673. [doi: 10.1007/s00134-018-5384-5] [Medline: 30259073]
6. Dziadzko MA, Harrison AM, Tiong IC, Pickering BW, Moreno Franco P, Herasevich V. Testing modes of computerized sepsis alert notification delivery systems. BMC Med Inform Decis Mak 2016;16(1):156 [FREE Full text] [doi: 10.1186/s12911-016-0396-y] [Medline: 27938401]
7. Meyer A, Zverinski D, Pfahringer B, Kempfert J, Kuehne T, Sündermann SH, et al. Machine learning for real-time prediction of complications in critical care: a retrospective study. Lancet Respir Med 2018;6(12):905-914. [doi: 10.1016/S2213-2600(18)30300-X] [Medline: 30274956]
8. Bhatia M, Sood SK. Temporal informative analysis in smart-ICU monitoring: m-healthcare perspective. J Med Syst 2016;40(8):190. [doi: 10.1007/s10916-016-0547-9] [Medline: 27388507]
9. Yamada T, Vacas S, Gricourt Y, Cannesson M. Improving perioperative outcomes through minimally invasive and non-invasive hemodynamic monitoring techniques. Front Med (Lausanne) 2018;5:144 [FREE Full text] [doi: 10.3389/fmed.2018.00144] [Medline: 29868596]
10. Poncette AS, Meske C, Mosch L, Balzer F. How to overcome barriers for the implementation of new information technologies in intensive care medicine. In: Proceedings of the 21st Human-Computer Interaction International Conference. 2019 Presented at: HCII '19; July 26-31, 2019; Orlando, FL p. 534-546. [doi: 10.1007/978-3-030-22649-7_43]
11. Baig MM, GholamHosseini H, Moqeem AA, Mirza F, Lindén M. A systematic review of wearable patient monitoring systems - current challenges and opportunities for clinical adoption. J Med Syst 2017;41(7):115. [doi: 10.1007/s10916-017-0760-1] [Medline: 28631139]
12. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
13. Tscholl DW, Handschin L, Rössler J, Weiss M, Spahn DR, Nöthiger CB. It's not you, it's the design - common problems with patient monitoring reported by anesthesiologists: a mixed qualitative and quantitative study. BMC Anesthesiol 2019;19(1):87 [FREE Full text] [doi: 10.1186/s12871-019-0757-z] [Medline: 31138143]
14. von Dincklage F, Suchodolski K, Lichtner G, Friesdorf W, Podtschaske B, Ragaller M. Investigation of the usability of computerized critical care information systems in Germany. J Intensive Care Med 2019;34(3):227-237. [doi: 10.1177/0885066617696848] [Medline: 28292221]
15. Wade VA, Eliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. Qual Health Res 2014;24(5):682-694. [doi: 10.1177/1049732314528809] [Medline: 24685708]
16. Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. JAMA 2018;319(12):1276-1278 [FREE Full text] [doi: 10.1001/jama.2018.1171] [Medline: 29584833]
17. Fairbanks RJ, Caplan S. Poor interface design and lack of usability testing facilitate medical error. Jt Comm J Qual Saf 2004;30(10):579-584. [doi: 10.1016/s1549-3741(04)30068-7] [Medline: 15518362]
18. 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;20(e1):e2-e8 [FREE Full text] [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]

19. Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 2005;74(7-8):519-526. [doi: 10.1016/j.ijmedinf.2005.01.003] [Medline: 16043081]
20. Bitkina OV, Kim HK, Park J. Usability and user experience of medical devices: an overview of the current state, analysis methodologies, and future challenges. Int J Ind Ergon 2020;76:102932. [doi: 10.1016/j.ergon.2020.102932]
21. Saeed N, Manzoor M, Khosravi P. An exploration of usability issues in telecare monitoring systems and possible solutions: a systematic literature review. Disabil Rehabil Assist Technol 2020;15(3):271-281. [doi: 10.1080/17483107.2019.1578998] [Medline: 30794009]
22. Yen PY, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc 2012;19(3):413-422 [FREE Full text] [doi: 10.1136/amiajnl-2010-000020] [Medline: 21828224]
23. Vincent CJ, Li Y, Blandford A. Integration of human factors and ergonomics during medical device design and development: it's all about communication. Appl Ergon 2014;45(3):413-419. [doi: 10.1016/j.apergo.2013.05.009] [Medline: 23778022]
24. Shah SG, Robinson I. Benefits of and barriers to involving users in medical device technology development and evaluation. Int J Technol Assess Health Care 2007;23(1):131-137. [doi: 10.1017/S0266462307051677] [Medline: 17234027]
25. Peischl B, Ferk M, Holzinger A. The fine art of user-centered software development. Software Qual J 2015;23(3):509-536. [doi: 10.1007/s11219-014-9239-1]
26. Roberts JP, Fisher TR, Trowbridge MJ, Bent C. A design thinking framework for healthcare management and innovation. Healthc (Amst) 2016;4(1):11-14. [doi: 10.1016/j.hjdsi.2015.12.002] [Medline: 27001093]
27. Johnson CM, Johnson TR, Zhang J. A user-centered framework for redesigning health care interfaces. J Biomed Inform 2005;38(1):75-87 [FREE Full text] [doi: 10.1016/j.jbi.2004.11.005] [Medline: 15694887]
28. Fidler R, Bond R, Finlay D, Guldenring D, Gallagher A, Pelter M, et al. Human factors approach to evaluate the user interface of physiologic monitoring. J Electrocardiol 2015;48(6):982-987. [doi: 10.1016/j.jelectrocard.2015.08.032] [Medline: 26375330]
29. Wiggermann N, Rempel K, Zerhusen RM, Pelo T, Mann N. Human-centered design process for a hospital bed: promoting patient safety and ease of use. Ergon Des 2019;27(2):4-12. [doi: 10.1177/1064804618805570]
30. Applying human factors and usability engineering to medical devices: guidance for industry and Food and Drug Administration staff. U.S. Food and Drug Administration. 2016. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices [accessed 2022-03-03]
31. Alexander G, Staggers N. A systematic review of the designs of clinical technology: findings and recommendations for future research. ANS Adv Nurs Sci 2009;32(3):252-279 [FREE Full text] [doi: 10.1097/ANS.0b013e3181b0d737] [Medline: 19707093]
32. Bazzano AN, Martin J, Hicks E, Faughnan M, Murphy L. Human-centred design in global health: a scoping review of applications and contexts. PLoS One 2017;12(11):e0186744 [FREE Full text] [doi: 10.1371/journal.pone.0186744] [Medline: 29091935]
33. Schmettow M, Schnittker R, Schraagen JM. An extended protocol for usability validation of medical devices: research design and reference model. J Biomed Inform 2017;69:99-114 [FREE Full text] [doi: 10.1016/j.jbi.2017.03.010] [Medline: 28336478]
34. Daniels J, Fels S, Kushniruk A, Lim J, Ansermino JM. A framework for evaluating usability of clinical monitoring technology. J Clin Monit Comput 2007;21(5):323-330. [doi: 10.1007/s10877-007-9091-y] [Medline: 17701381]
35. Gravetter FJ, Forzanno LB. Research methods for the behavioral sciences. 4th edition. Belmont, CA: Wadsworth Publishing; 2012.
36. Quick guide for clinician users: vital SyncTM virtual patient monitoring platform. Medtronic. 2017. URL: https://tinyurl.com/44tjfek9 [accessed 2022-03-03]
37. Vital SyncTM virtual patient monitoring platform: user guide. Medtronic. 2017. URL: https://tinyurl.com/2p8mcp52 [accessed 2022-03-03]
38. Health informatics and monitoring: vital SyncTM virtual patient monitoring platform. Medtronic. URL: https://tinyurl.com/5cbzu9ts [accessed 2019-01-15]
39. Ergonomie der Mensch-System-Interaktion - Methoden zur Gewährleistung der Gebrauchstauglichkeit, die eine benutzer-orientierte Gestaltung unterstützen. Beuth Publishing DIN. 2002. URL: https://www.beuth.de/de/technische-regel/iso-tr-16982/57203361 [accessed 2022-03-03]
40. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. Qual Health Res 1993;3(4):430-441. [doi: 10.1177/104973239300300403]
41. Wiklund ME, Kendler J, Strochlic AY. Selecting tasks. In: Wilkund ME, Kendler J, Strochlic AY, editors. Usability testing of medical devices. 2nd edition. Boca Raton, FL: CRC Press; 2016:261-316.
42. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
43. Lewis JR. The system usability scale: past, present, and future. Int J Hum-Comput Interact 2018;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]

44. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004;37(1):56-76 [FREE Full text] [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
45. Sauro J, Lewis J. Quantifying user research. In: Sauro J, Lewis J, editors. Quantifying the user experience: practical statistics of user research. Cambridge, MA: Morgan Kaufmann; 2016:9-18.
46. Rubin J, Chisnell D, Spool J. Handbook of usability testing: how to plan, design, and conduct effective tests. 2nd edition. Indianapolis, IN: Wiley; 2008.
47. Broekhuis M, van Velsen L, Hermens H. Assessing usability of eHealth technology: a comparison of usability benchmarking instruments. Int J Med Inform 2019;128:24-31. [doi: 10.1016/j.ijmedinf.2019.05.001] [Medline: 31160008]
48. Nielsen J, Budiu R. Success rate: the simplest usability metric. Nielsen Norman Group. 2001. URL: https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/ [accessed 2019-11-06]
49. Schnittker R, Schmettow M, Verhoeven F, Schraagen JM. Combining situated Cognitive Engineering with a novel testing method in a case study comparing two infusion pump interfaces. Appl Ergon 2016;55:16-26. [doi: 10.1016/j.apergo.2016.01.004] [Medline: 26995032]
50. Albert B, Tullis T. Measuring the user experience: collecting, analyzing, and presenting usability metrics. 2nd edition. Amsterdam, The Netherlands: Morgan Kaufmann; 2013.
51. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52(3/4):591-611. [doi: 10.2307/2333709]
52. Levene H. Robust tests for equality of variances. In: Olkin I, editor. Contributions to probability and statistics: essays in honor of Harold hotelling. Redwood City, CA: Stanford University Press; 1960:278-292.
53. Poncette AS, Mosch L, Spies C, Schmieding M, Schiefenhövel F, Krampe H, et al. Improvements in patient monitoring in the intensive care unit: survey study. J Med Internet Res 2020;22(6):e19091 [FREE Full text] [doi: 10.2196/19091] [Medline: 32459655]
54. Poncette AS, Spies C, Mosch L, Schieler M, Weber-Carstens S, Krampe H, et al. Clinical requirements of future patient monitoring in the intensive care unit: qualitative study. JMIR Med Inform 2019;7(2):e13064 [FREE Full text] [doi: 10.2196/13064] [Medline: 31038467]
55. Turner P, Kushniruk A, Nohr C. Are we there yet? Human factors knowledge and health information technology - the challenges of implementation and impact. Yearb Med Inform 2017;26(1):84-91 [FREE Full text] [doi: 10.15265/IY-2017-014] [Medline: 29063542]
56. Labeling: regulatory requirements for medical devices. U.S. Food & Drug Administration. 1997. URL: https://www.fda.gov/files/medical%20devices/published/Labeling---Regulatory-Requirements-for-Medical-Devices-(FDA-89-4203).pdf [accessed 2022-03-03]
57. Roman LC, Ancker JS, Johnson SB, Senathirajah Y. Navigation in the electronic health record: a review of the safety and usability literature. J Biomed Inform 2017;67:69-79 [FREE Full text] [doi: 10.1016/j.jbi.2017.01.005] [Medline: 28088527]
58. Moradian S, Krzyzanowska MK, Maguire R, Morita PP, Kukreti V, Avery J, et al. Usability evaluation of a mobile phone-based system for remote monitoring and management of chemotherapy-related side effects in cancer patients: mixed-methods study. JMIR Cancer 2018;4(2):e10932 [FREE Full text] [doi: 10.2196/10932] [Medline: 30578238]
59. Andrade E, Quinlan L, Harte R, Byrne D, Fallon E, Kelly M, et al. Novel interface designs for patient monitoring applications in critical care medicine: human factors review. JMIR Hum Factors 2020;7(3):e15052 [FREE Full text] [doi: 10.2196/15052] [Medline: 32618574]
60. Flohr L, Beaudry S, Johnson KT, West N, Burns CM, Ansermino JM, et al. Clinician-driven design of VitalPAD - an intelligent monitoring and communication device to improve patient safety in the intensive care unit. IEEE J Transl Eng Health Med 2018;6:3000114 [FREE Full text] [doi: 10.1109/JTEHM.2018.2812162] [Medline: 29552425]
61. Faiola A, Srinivas P, Hillier S. Improving patient safety: integrating data visualization and communication into ICU workflow to reduce cognitive load. Proc Int Symp Hum Factors Ergon Healthc 2015;4(1):55-61. [doi: 10.1177/2327857915041013]
62. Faiola A, Srinivas P, Duke J. Supporting clinical cognition: a human-centered approach to a novel ICU information visualization dashboard. AMIA Annu Symp Proc 2015;2015:560-569 [FREE Full text] [Medline: 26958190]
63. Chuang CH, Tseng PC, Lin CY, Lin KH, Chen YY. Burnout in the intensive care unit professionals: a systematic review. Medicine (Baltimore) 2016;95(50):e5629 [FREE Full text] [doi: 10.1097/MD.0000000000005629] [Medline: 27977605]
64. Richardsen AM, Martinussen M, Kaiser S. Stress, human errors and accidents. In: Burke RJ, Richardsen AM, editors. Increasing occupational health and safety workplaces: individual, work and organizational factors. Cheltenham, UK: Edward Elgar Publishing; 2019:48-67.
65. Kumar A, Pore P, Gupta S, Wani AO. Level of stress and its determinants among intensive care unit staff. Indian J Occup Environ Med 2016;20(3):129-132 [FREE Full text] [doi: 10.4103/0019-5278.203137] [Medline: 28446837]
66. Graber ML, Kissam S, Payne VL, Meyer AN, Sorensen A, Lenfestey N, et al. Cognitive interventions to reduce diagnostic error: a narrative review. BMJ Qual Saf 2012;21(7):535-557. [doi: 10.1136/bmjqs-2011-000149] [Medline: 22543420]
67. Kahn JM. What we talk about when we talk about intensive care unit strain. Ann Am Thorac Soc 2014;11(2):219-220. [doi: 10.1513/AnnalsATS.201311-406ED] [Medline: 24575987]

68. The Topol review: preparing the healthcare workforce to deliver the digital future: an independent report on behalf of the Secretary of State for Health and Social Care. National Health Service. 2019. URL: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf [accessed 2022-03-03]
69. Gonzalez Velasco JM. Complex and transdisciplinary strategies for promotion and prevention in digital health: towards the ecology of knowledge. In: Mantel-Teeuwisse A, Khatri B, Uzman N, Mellianti S, editors. FIP Digital health in pharmacy education: developing a digitally enabled pharmaceutical workforce. The Hague, The Netherlands: International Pharmaceutical Federation; 2021:106-108.
70. Machleid F, Kaczmarczyk R, Johann D, Balčiūnas J, Atienza-Carbonell B, von Maltzahn F, et al. Perceptions of digital health education among European medical students: mixed methods survey. J Med Internet Res 2020;22(8):e19827 [FREE Full text] [doi: 10.2196/19827] [Medline: 32667899]
71. Matern U. Design, usability and staff training - what is more important? In: Doffy VG, editor. Advances in human aspects of healthcare. Boca Raton, FL: CRC Press; 2012:426-428.
72. Poncette AS, Glauert DL, Mosch L, Braune K, Balzer F, Back DA. Undergraduate medical competencies in digital health and curricular module development: mixed methods study. J Med Internet Res 2020;22(10):e22161 [FREE Full text] [doi: 10.2196/22161] [Medline: 33118935]
73. Holeman I, Kane D. Human-centered design for global health equity. Inf Technol Dev 2019;26(3):477-505 [FREE Full text] [doi: 10.1080/02681102.2019.1667289] [Medline: 32982007]
74. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
75. Laboratories: human factors engineering: Susan Hallbeck. Mayo Clinic. URL: https://www.mayo.edu/research/labs/human-factors-engineering/overview [accessed 2020-05-09]

Abbreviations

HCD: human-centered design
ICU: intensive care unit
IT: information technology
SUS: System Usability Scale
UI: user interface

Edited by G Eysenbach; submitted 25.05.21; peer-reviewed by N West; comments to author 05.07.21; revised version received 13.08.21; accepted 19.09.21; published 11.03.22.

Please cite as:
Poncette AS, Mosch LK, Stablo L, Spies C, Schieler M, Weber-Carstens S, Feufel MA, Balzer F
A Remote Patient-Monitoring System for Intensive Care Medicine: Mixed Methods Human-Centered Design and Usability Evaluation
JMIR Hum Factors 2022;9(1):e30655
URL: https://humanfactors.jmir.org/2022/1/e30655
doi:10.2196/30655
PMID:35275071

Original Paper

Development of a Digital Support Application With Evidence-Based Content for Sustainable Return to Work for Persons With Chronic Pain and Their Employers: User-Centered Agile Design Approach

Christina Turesson[1*], BSc OT, PhD; Gunilla Liedberg[1*], MSc OT, PhD; Mathilda Björk[2*], BSc OT, PhD

[1]Division of Prevention, Rehabilitation and Community Medicine, Department of Health, Medicine and Caring Sciences, Linköping University, Norrköping, Sweden

[2]Pain and Rehabilitation Centre, Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

[*]all authors contributed equally

Corresponding Author:
Christina Turesson, BSc OT, PhD
Division of Prevention, Rehabilitation and Community Medicine
Department of Health, Medicine and Caring Sciences
Linköping University
Kåkenhus plan 7
Norrköping, 601 74
Sweden
Phone: 46 11363602
Email: christina.turesson@liu.se

Abstract

Background: Persons with chronic pain experience a lack of support after completing rehabilitation and the responsibility for the return-to-work (RTW) process is taken over by the employer. In addition, employers describe not knowing how to support their employees. Smartphone apps have been increasingly used for self-management, but there is a lack of available eHealth apps with evidence-based content providing digital support for persons with chronic pain and their employers when they return to work.

Objective: This study aims to describe the development of a digital support application with evidence-based content that includes a biopsychosocial perspective on chronic pain for sustainable RTW for persons with chronic pain and their employers (SWEPPE [Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers]).

Methods: A user-centered agile design approach was applied. The multidisciplinary project team consisted of health care researchers, a user representative, and a software team. A total of 2 reference groups of 7 persons with chronic pain and 4 employers participated in the development process and usability testing. Mixed methods were used for data collection. The design was revised using feedback from the reference groups. The content of SWEPPE was developed based on existing evidence and input from the reference groups.

Results: The reference groups identified the following as important characteristics to include in SWEPPE: keeping users motivated, tracking health status and work situation, and following progress. SWEPPE was developed as a smartphone app for the persons with chronic pain and as a web application for their employers. SWEPPE consists of six modules: the action plan, daily self-rating, self-monitoring graphs, the coach, the library, and shared information with the employer. The employers found the following functions in SWEPPE to be the most useful: employees' goals related to RTW, barriers to RTW, support wanted from the employer, and the ability to follow employees' progress. The persons with chronic pain found the following functions in SWEPPE to be the most useful: setting a goal related to RTW, identifying barriers and strategies, and self-monitoring. Usability testing revealed that SWEPPE was safe, useful (ie, provided relevant information), logical, and easy to use with an appealing interface.

Conclusions: This study reports the development of a digital support application for persons with chronic pain and their employers. SWEPPE fulfilled the need of support after an interdisciplinary pain rehabilitation program with useful functions such as setting a goal related to RTW, identification of barriers and strategies for RTW, self-monitoring, and sharing information between the employee and the employer. The user-centered agile design approach contributed to creating SWEPPE as a relevant and easy-to-use eHealth intervention. Further studies are needed to examine the effectiveness of SWEPPE in a clinical setting.

(JMIR Hum Factors 2022;9(1):e33571) doi:10.2196/33571

KEYWORDS

agile design process; chronic pain; digital support; eHealth; return to work; self-management; smartphone apps; user-centered design; mobile phone

Introduction

Background

Chronic musculoskeletal pain, which affects 10% to 20% of the European population and negatively impacts functioning, quality of life, and the ability to work, comes with significant individual and societal expenses, including costs associated with sick leave and loss of productivity [1-6]. A recently published interview study [7] showed that persons with chronic pain experienced a lack of support after completing a rehabilitation program when responsibility for the return-to-work (RTW) process was taken over by the employer. In addition, the employers reported lacking knowledge on how to support their employees' RTW and requested more knowledge about how chronic pain might affect work status and the needs and challenges their employees with chronic pain might experience [7]. For a successful RTW, employers need to effectively collaborate, communicate, and negotiate with their employees, all interactions that require good listening skills [8,9]. In general, barriers to RTW for persons with chronic pain include lack of workplace support, lack of relationships with supervisors and coworkers, and inability to find the right fit between a person's physical abilities and job tasks [10,11]. A smartphone app could be used to deal with the above challenges, improve the rehabilitation process, and counteract passivity by increasing interaction between the employer and the employee [12], leading to a shared decision-making model [13] used in work rehabilitation to increase a successful outcome in the RTW process.

Digital support (web-based applications or smartphone apps) is a growing intervention for persons with chronic pain and a useful tool for quality of learning [14-16]. The strengths of digital interventions include evidence-based content; possibility for daily registrations of health aspects; simple design with short, easily readable texts [17]; and reading about other people's experiences [18]. Typically, self-management includes providing knowledge and education about the condition (including its consequences) and self-assessment of health [17,19,20]. This can contribute to the individuals' learning about their own capacity [21,22], which can lead to an increased sense of control and motivation for continued self-management [23].

Digital applications can be valuable tools for persons with chronic pain, especially when used in an outclinic setting [24], and can reduce pain and disability [25,26]. Despite these positive effects, research has reported limitations related to the low overall quality of smartphone apps for chronic pain and the lack of rigorous assessment of their effectiveness [27,28]. Therefore, combining evidence-based concepts with stakeholder involvement in the development of eHealth interventions is highly important [15,27]. The key elements of user-centered design (UCD) approaches include stakeholder involvement, iterative design and development, user stories, user personas, interviews, prototyping, and usability testing to identify and fulfill the users' needs and requirements [29-33]. To manage challenges such as incorrect clinical or user context or flaws in evaluation [34], researchers need to use a multidisciplinary development approach, continuous and systematic evaluation, and robust evaluation methods [35].

Objectives

Web-based support for RTW has shown to be successful and cost-effective for persons with musculoskeletal disorders [36]. However, to the best of our knowledge, no evidence-based digital support exists that improves sustainable RTW for persons with chronic pain and their employers. To fill this gap in knowledge, the aim of this study is to develop a digital support application with evidence-based content that includes a biopsychosocial perspective on chronic pain for sustainable RTW for persons with chronic pain and their employers: SWEPPE (Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers).

Methods

Study Design

In this study, a user-centered agile design [30] was used. Five principles guided the process [30]: (1) separate product discovery and product creation phases; (2) iterative design and development with empirical feedback to revise designs in the next step; (3) parallel design and development activities using one sprint ahead; (4) continuous involvement of users via reference groups; and (5) artifact-mediated communication via user personas and scenarios (Figure 1).

The multidisciplinary project team consisted of health care researchers, a user representative, and a software team (Table 1).

Figure 1. Flowchart of practices and data collection during the development process to create SWEPPE (Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers). Principles guiding the process: (1) separate product discovery and product creation phases; (2) iterative design and development with empirical feedback to revise designs in the next step; (3) parallel design and development activities using one sprint ahead; (4) continuous involvement of users via reference groups through the process; (5) artifact-mediated communication via user personas and scenarios. SUS: System Usability Scale.

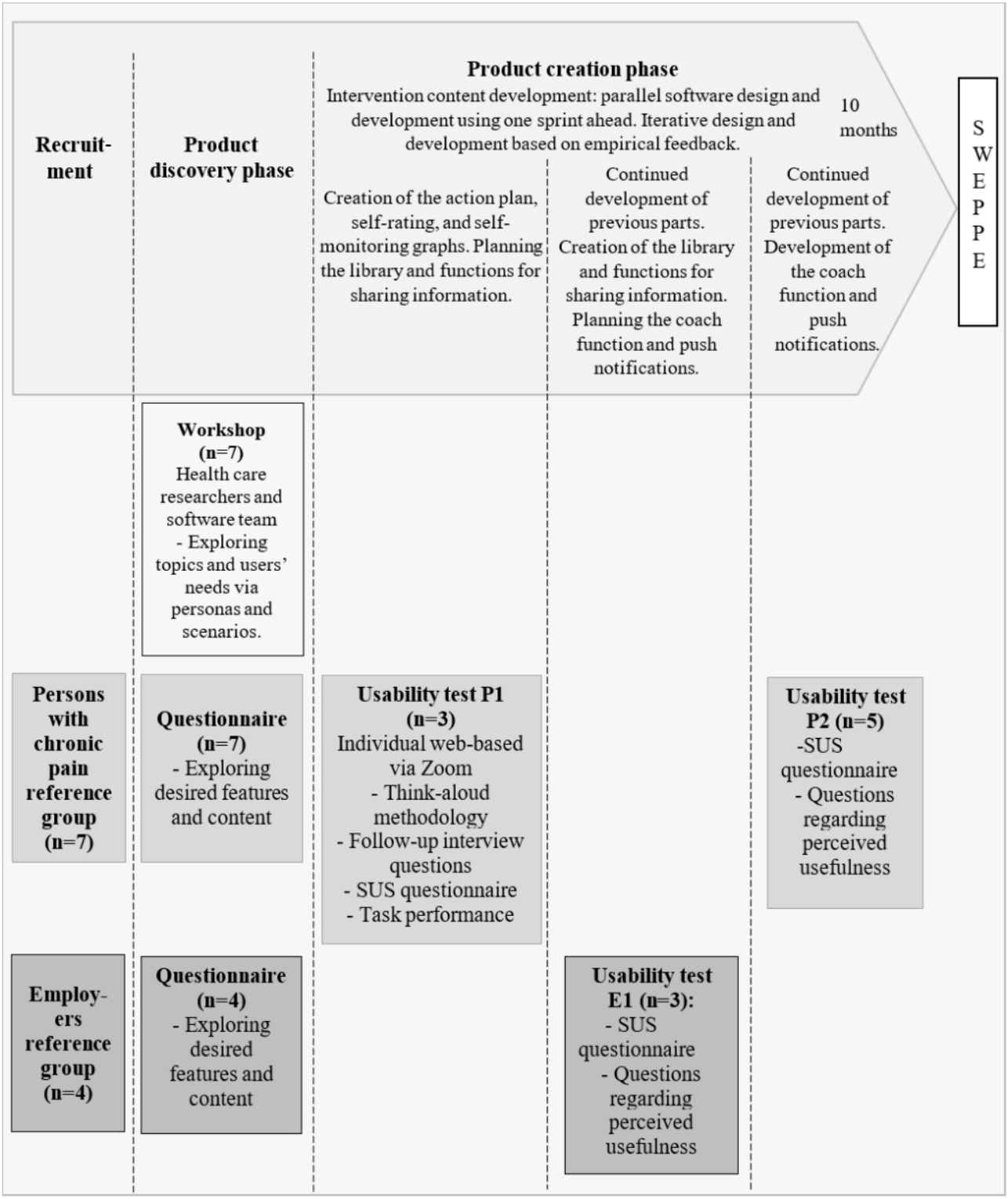


Table 1. Description of the multidisciplinary project team.

Grouping	Total, n (%)	Pain expertise, n (%)	Return-to-work expertise, n (%)	Licensed health care providers, n (%)	Electronic apps expertise, n (%)
Health care researchers[a]	3 (100)	2 (67)	2 (67)	3 (100)	0 (0)
User representative[b]	1 (100)	1 (100)	0 (0)	0 (0)	0 (0)
Software team[c]	5 (100)	0 (0)	0 (0)	0 (0)	5 (100)

[a]PhD occupational therapist.

[b]A research partner from the Swedish Rheumatism Association.

[c]Including user experience design, back-end and front-end development, and project management.

The development process was led by a senior researcher (MB) with extensive experience with RTW, chronic pain, and occupational therapy. Two reference groups representing the end users—patients with chronic pain and their employers—were recruited for the development process. The initial product discovery phase consisted of exploring the users' needs and wishes about the functions and contents of SWEPPE. The second product creation phase involved design, development, and usability testing of SWEPPE. Mixed methods were applied to collect quantitative and qualitative data for early input and feedback from the users throughout the process.

Ethics Approval

The Swedish Ethical Review Board approved the study (Dnr 2020-01593).

Participants in the Reference Groups

Participants in the reference groups were recruited using a relevance sampling strategy [37] conducted at a pain and rehabilitation clinic in southern Sweden. For persons with chronic pain, the following inclusion criteria were used: employed, participation in an interdisciplinary pain rehabilitation program (IPRP) within the last 2 years, and interest in participating in the development of the application. For the employers, the following inclusion criteria were used: experience with an employee who had chronic pain and interest in participating in the development of the application. A total of 12 persons with chronic pain met the inclusion criteria and were invited to the study by email. The invitation contained information about the study, and a telephone follow-up conversation was conducted after approximately 1 week. Of these 12 persons, 4 (33%) did not respond to the invitation or follow-up call, and 1 (8%) declined participation. Thus, of the 12 persons invited, 7 (58%)—4 women and 3 men—with a mean age of 45 (SD 9; range 36-60) years provided informed consent and were included in the study (Table 2).

Table 2. Background data on the persons with chronic pain participating in the reference group (n=7).

Characteristic	Values, n (%)
Reported years living with pain	
0-7	3 (43)
8-14	3 (43)
>15	1 (14)
Types of pain	
Back or neck pain	4 (57)
Nerve pain or neuropathic pain	1 (14)
Fibromyalgia	2 (29)
Other[a]	3 (43)
Employment status	
Working or studying full-time	4 (57)
Working or studying part-time	2 (29)
Sick leave	1 (14)

[a]Leg pain, migraine, and Horton disease.

The mean time since participation in IPRP was 10 (SD 5; range 4-19) months. Of the 7 participants, 1 (14%) had experienced a large degree of support during RTW after IPRP, and 6 (86%) had experienced some support from different stakeholders (employer, health care, or the social insurance agency). A total of 10 employers who previously had been involved in an RTW process for an employee with chronic pain at the rehabilitation clinic were invited to the study in a similar way as the persons with chronic pain. Of these 10 employers, 3 (30%) were not possible to reach and 3 (30%) declined participation in the study

owing to lack of time. Finally, of the 10 employers invited, the 4 (40%)—3 men and 1 woman—who agreed to participate in the study were from private and public areas of the labor market (school and education, services and sales, building and manufacturing, or machine and transportation).

Practices to Develop SWEPPE

Overview

Practices such as workshop and usability testing were applied during the different phases of the development of SWEPPE (Figure 1). The reference group of persons with chronic pain participated in 2 usability tests, and the employers participated in 1.

Workshop

User personas and two scenarios (*using SWEPPE* and *being back at work*) were developed based on previous research and clinical expertise with persons with chronic pain and other stakeholders. The user personas and scenarios were used in a workshop where 2 health care researchers and the software team verbally and visually presented information about different types of users [38] and about how to bring the needs of the persons with chronic pain and their employers into the development process. For each scenario, brainstorming was performed about what a user persona might think, feel, do, and say in a given situation.

Usability Testing

The first usability test was performed with the reference groups of persons with chronic pain (Figure 1, P1) and consisted of both formative and summative testing [39], where screen layouts with partial functionality were used. The test sessions were conducted on the web via Zoom (Zoom Video Communications) meetings owing to the COVID-19 pandemic. Before the first usability test, a pilot test conducted with a person not involved in the project was performed, which led to minor adjustments of the test situation. Then, 3 participants from the reference group of persons with chronic pain participated in individual usability testing sessions. The tests were led by the user experience designer, and the first author (CT) participated as an observer and took notes. During the tests, the participants were given tasks to perform and were asked to verbalize their experience—that is, a think-aloud methodology [40] was applied. All meetings were recorded and shared digitally with 3 members of the software team, who also took notes as part of the think-aloud methodology to aid the development process.

The usability test with the reference group of employers (Figure 1, E1) was performed on a functioning web application. Then, 2 fictional *employees* were created based on the user personas used in the workshop. In addition, 2 researchers acting as these fictional *employees* created accounts in the SWEPPE smartphone app and invited their *employers* in the reference group to access the web application on their own computer. In the web application, the employers could access the library and follow the goals and self-reported data of their employees. The employers received updated information regarding progress from the employees for 3 weeks.

The second usability test with the reference group of persons with chronic pain (Figure 1, P2) was performed on a functioning smartphone app. The persons with chronic pain downloaded and tested SWEPPE at home on their own smartphone for 2 weeks.

Data Collection

Overview

Data were collected from the workshop, digital questionnaires, and usability testing during the development of SWEPPE (Figure 1).

Workshop

The workshop generated 2 empathy maps [41] that included short statements of what the user personas might think, feel, do, and say in the given scenarios. These maps were used for the identification of topics, questions, or needs to be considered while developing SWEPPE.

Questionnaire

The questionnaire focused on exploring the participants' experiences of using smartphone apps and suggestions for the functions and content of SWEPPE. The participants in the reference groups were asked to rate the importance of the different proposed functions on a 10-point scale ranging from 1 (not important) to 10 (very important). The proposed functions were based on aspects identified as strengths in previous research, such as setting a goal related to RTW [42-44], the possibility to monitor health status [12], access to a knowledge base about pain and positive examples of RTW [18], digital coaching, and access to frequently asked questions or stories of persons with chronic pain [18,36,42,45].

Usability Testing

During the web-based usability test with the persons with chronic pain (P1), data were collected in several steps. First, notes were taken by the first author (CT) and 2 or 3 members of the project team during the think-aloud process regarding what the person said and did when performing the assignments in the SWEPPE prototype. If the test persons were silent, they were prompted by the test leader with questions such as "What are you thinking right now?" Second, at the end of each usability test, the participants were asked open-ended questions regarding the overall impression of SWEPPE: positive or negative functions and content of SWEPPE, what was missing or could be improved, how they would describe SWEPPE to a colleague or friend, and their opinion on how the SWEPPE prototype would be helpful in RTW. Third, after the usability tests, the participants were asked to fill out the System Usability Scale (SUS) questionnaire [46,47] for the global usability assessment of the SWEPPE prototype. The SUS consists of 10 items rated on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). A total SUS score was calculated, ranging from 0 to 100 (higher scores represent better usability) [46]. A total SUS score >70 represents good usability [48]. Fourth, task performance [39] was assessed during the first usability test where the number of correctly completed tasks was registered by the first author (CT).

After usability tests P2 and E1, both reference groups received the SUS questionnaire and follow-up questions regarding the functions in SWEPPE. The participants were asked to rate the perceived usefulness of different parts of SWEPPE as a support on a scale ranging from 0 (no support) to 100 (maximum support). They were also given the opportunity to comment on the functions and content of SWEPPE—for example, what was missing or could be improved and how they would describe SWEPPE to a colleague or friend.

Data Analysis

Quantitative data from the surveys and the SUS questionnaire were summarized and analyzed using descriptive statistics. The qualitative data gathered in the empathy maps were summarized and grouped based on topics, questions, and needs to address in the development of SWEPPE.

The notes collected during the think-aloud methodology and data from the open questions in the first usability test with the reference group of persons with chronic pain were analyzed using Instant data analysis [40]. Instant data analysis was performed after each of the usability test sessions, where the test leader, the observer (CT), and 3 members of the software team participated in a Zoom meeting to discuss their notes and the usability problems that had been identified. All usability issues were written down and sorted into groups based on the assignments performed in the prototype: create an account, set a goal related to RTW, review the action plan, finish settings, register daily health status, or follow progress in the overview. The identified usability problems were then discussed by the whole project team and were used to guide adjustments to the SWEPPE prototype before finalizing the application.

Results

Product Discovery Phase

The workshop with the project group generated topics, questions, or needs considered during the development process of SWEPPE (Table 3).

Results from the questionnaire about the desired content and functions of SWEPPE showed that maintaining motivation and following progress were of great importance for both reference groups (Tables 4 and 5).

Table 3. Identified topics, questions, or needs during the workshop with the project group and empathy mapping of user persona Carina for different scenarios and how these were addressed in SWEPPE[a].

Scenario and identified topics, questions, and needs to consider in the development process	Addressed in SWEPPE
Being back at work	
Will I manage? Do I have the skills needed?	Goal setting, self-monitoring, and overview to support insights about one's capacity.
Will I get the support I need from the employer?	Identify support wanted from the employer and possibility to share with the employer.
Manage balance between work and leisure.	Goal setting, self-monitoring, and overview for feedback.
Find a daily routine.	Self-monitoring and overview.
Learning new ways.	Library, self-monitoring, and overview.
Apply strategies learned during rehabilitation.	Identify barriers to RTW[b] and strategies to handle them; self-monitoring.
Using SWEPPE	
How will SWEPPE help me RTW?	Overarching question guiding the general design of all the functions in SWEPPE.
Using SWEPPE must be quick and easy.	General design of SWEPPE application as quick and easy to use and demanding low cognitive load.
Difficult at first when I started.	General design of SWEPPE when creating and setting up a new account.
Uncertain about what data the employer can see in SWEPPE.	General design of SWEPPE with easy access to information the user wants to share with the employer.
Feeling guilty if not using SWEPPE every day.	Are data presented in the overview in a useful way even if data are missing?
Proud and happy about her progress.	Design of overview for easy visualization of progress.

[a]SWEPPE: Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers.

[b]RTW: return to work.

Table 4. Results from the initial survey with persons with chronic pain (n=7) regarding the desired content and functions of SWEPPE[a].

Questions, desired content, and topics of interest	Rating of importance, median (IQR)
An application for people with chronic pain and their employer as support for return to work (SWEPPE) would be interesting	
To keep me motivated	10 (9-10)
To follow and focus on my progress	9 (7-10)
To keep track of my health status	8 (8-10)
To keep track of my work situation	8 (4-10)
To get inspiration from others	8 (59)
Important information to know about the application	
Security details or privacy information	10 (8-10)
How to optimize usability	8 (8-10)
Where information in the app comes from	8 (8-10)
Desired content or topics of information in SWEPPE	
Pain and coping with pain	10 (9-10)
Stress and coping	9 (8-10)
Work and work ability	9 (7-10)
Ergonomics	9 (6-10)
Thoughts and feelings	8 (8-9)
Balance in daily activities	8 (8-10)
Coping during hard times	8 (8-9)
Healthy lifestyle	8 (7-10)
Others' experiences of coping with chronic pain	8 (7-9)
Workplace adaptation	8 (8-9)
Communication, relations, social support	7 (6-9)
Desired functions in SWEPPE	
Setting goals	9 (8-10)
Communicate with a coach	9 (7-10)
FAQ (frequently asked questions) available	8 (7-9)
Tips on workplace adaptation	8 (7-9)
Communicate information with my employer	8 (7-9)
Important functions SWEPPE should have	
Push notifications	9 (2-10)
Adaptive functions	8 (7-9)
Adaptive design	8 (6-8)
Download information	6 (5-10)
Preferred health aspects to record in SWEPPE or receive information about from employee	
Pain	10 (10-10)
Sleep	10 (8-10)
Physical activity	10 (8-10)
Work situation	10 (6-10)
Balanced life situation	8 (7-10)
Workload	7 (5-9)

[a]SWEPPE: Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers.

Table 5. Results from the initial survey with employers (n=4) regarding the desired content and functions of SWEPPE[a].

Questions, desired content, and topics of interest	Rating of importance, median (IQR)
An application for people with chronic pain and their employer as support for RTW[b] (SWEPPE) would be interesting	
To motivate and support the employee	9.5 (9-10)
To follow the employee's progress	9.5 (9-10)
To receive information about my responsibility as an employer	9.5 (9-10)
To receive tips on adaptation of the work situation	9.5 (9-10)
To follow the employee's work situation	8.5 (8-9)
To follow the employee's health status	8.5 (8-9)
To receive information about chronic pain	7.5 (7-9)
To get inspiration from others	6.5 (6-8)
Important information to know about the application	
How to optimize usability	8.5 (8-9)
Where information in the app comes from	8.5 (8-9)
Security details or privacy information	8.5 (8-9)
Desired content or topics of information in SWEPPE	
Work and work ability	9.5 (9-10)
Ergonomics	9 (8-9)
Information about my responsibility as an employer	9 (9-10)
Stress and coping	9 (9-9)
About pain and coping with pain	9 (9-9)
Workplace adaptation	9 (8-9)
Balance in daily activities	8.5 (8-9)
Coping during hard times	8 (7-9)
Communication, relations, social support	8 (7-8)
Thoughts and feelings	8 (6-9)
Healthy lifestyle	7 (5-9)
Others' experiences of coping with chronic pain	6 (4-8)
Desired functions in SWEPPE	
Receive information about the employee's goals	9 (9-9)
Tips on workplace adaptation	9 (9-9)
FAQ[c] available	7 (6-9)
Important functions SWEPPE should have	
Adaptive design	8 (7-9)
Adaptive functions	6.5 (6-8)
Download information	6.5 (6-7)
Push notifications	4.5 (2-7)
Preferred health aspects to record in SWEPPE or receive information about from employee	
Work situation	10 (10-10)
Workload	10 (10-10)
Pain	9.5 (9-10)
Physical activity	9 (9-9)
Sleep	8.5 (8-9)

Questions, desired content, and topics of interest	Rating of importance, median (IQR)
Balanced life situation	8.5 (7-9)

[a]SWEPPE: Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers.

[b]RTW: return to work.

[C]FAQ: frequently asked questions.

For the persons with chronic pain, the opportunity to keep track of their health status and work situation was also important. Employers wanted information about their responsibility and suggestions for adapting the work situation. Both reference groups preferred getting feedback through graphs showing changes over time. Persons with chronic pain wanted to use SWEPPE on their smartphone, and most (4/7, 57%) reported wanting to use SWEPPE daily. The employers had a diverse view of how often they would use SWEPPE. An employer wanted to use SWEPPE when needed, another weekly, and another monthly. Most of the persons with chronic pain preferred recording pain (5/7, 71%) and sleep (4/7, 57%) daily and physical activity (4/7, 57%), balanced life situation (5/7, 71%), and work situation (4/7, 57%) weekly. The opinion among the persons with chronic pain about push notifications was mixed: push notifications were rated as an important function (Tables 4 and 5), but a majority (4/7, 57%) did not want them included in the SWEPPE application. However, most participants noted that their acceptance of push notifications would depend on the available settings. The other characteristics that the reference groups rated as important were compatibility with a smartphone and ease of use. The results from the questionnaire were used to prioritize the functions and development of SWEPPE.

Product Creation Phase

The initial development of SWEPPE, based on the first survey of persons with chronic pain, focused on three aspects: (1) the action plan, where the users assess their work ability, set a goal related to RTW, identify barriers to RTW, develop strategies to handle barriers, and identify support wanted from the employer; (2) self-rating, where the users register daily health status and work situation; and (3) self-monitoring graphs, where the users follow their progress and receive feedback to keep them motivated.

These 3 aspects were developed and tested along with the overall design (eg, colors and layout) in the first usability test (P1). The participants in usability test P1 were in general positive to the prototype and experienced it as relevant, quick, and easy to use. They stressed the importance of SWEPPE not demanding too much of them cognitively. They described SWEPPE as a tool to help them stay motivated and learn more about themselves and their pain. The task performance rate was high (Table 6).

Some usability problems were identified using the think-aloud methodology and were addressed in the continued development process (Figure 2).

Table 6. Results from usability testing P1 and P2 with the persons with chronic pain and E1 with the employers. Data collection from assessment of task performance and questionnaires.

Time points	Usability test: P1 persons with chronic pain (n=3)	Usability test: E1 employers (n=3)	Usability test: P2 persons with chronic pain (n=6)
Task performance[a], n (%)			
Create an account	3 (100)	N/A[b]	N/A
Set a goal	3 (100)	N/A	N/A
Review the action plan	3 (100)	N/A	N/A
Finish action plan settings	3 (100)	N/A	N/A
Register daily health status	3 (100)	N/A	N/A
Follow progress in the overview	3 (100)	N/A	N/A
SUS[c] questionnaire, median (IQR)			
SUS score point[d]	95 (94-98)	88 (72-89)	86.5 (77-94)
Employers perceived usefulness[e] of receiving information, median (IQR)			
About the employee's work-related goal	N/A	74 (58.5-83.5)	N/A
About barriers for RTW[f] identified by the employee	N/A	71 (61.5-85.5)	N/A
About strategies identified by the employee	N/A	46 (32.5-59.5)	N/A
About support wanted from the employer	N/A	73 (67.5-86.5)	N/A
To follow the employee's progress in a graph	N/A	74 (70.5-87)	N/A
From the library	N/A	50 (31-62)	N/A
To be reminded of using SWEPPE[g]	N/A	100 (55-100)	N/A
Persons with chronic pains perceived usefulness of SWEPPE, median (IQR)			
Setting a work-related goal and following the progress	N/A	N/A	81 (53.3-92.3)
Identifying barriers and strategies for RTW	N/A	N/A	68 (53-90.5)
Sharing information with the employer	N/A	N/A	53.5 (28.3-60.8)
Self-monitoring health aspects and getting an overview	N/A	N/A	80 (56-88.3)
Using the library	N/A	N/A	60.5 (54-75.3)
Asking questions and receiving answers from the coach	N/A	N/A	47 (41.5-69)
Getting reminders of daily self-rating of health aspects and weekly evaluation of goal fulfillment	N/A	N/A	85.5 (70.8-95.8)

[a]Number of correctly completed tasks.

[b]N/A: not applicable.

[c]SUS: System Usability Scale.

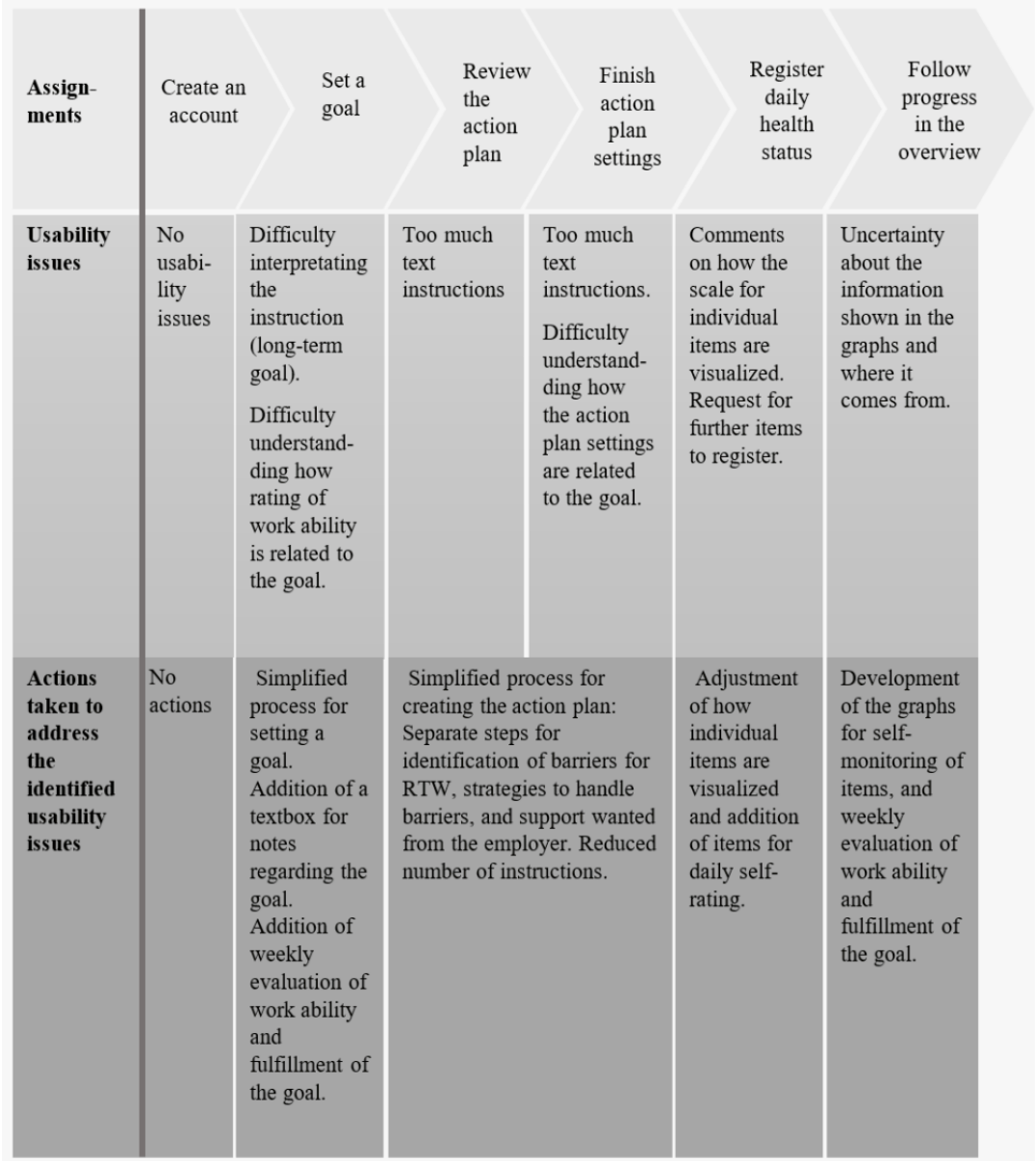
[d]SUS score points range from 0 to 100, where higher scores represent better usability.

[e]Rated on a visual analogue scale ranging from 0 (no support) to 100 (maximum support).

[f]RTW: return to work.

[g]SWEPPE: Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers.

Figure 2. Overview of the assignments in usability test P1 performed by participants (n=3) from the reference group of persons with chronic pain, the usability issues identified during the think-aloud methodology, and how these issues were addressed in the continued development process. RTW: return to work.



The process of setting a goal related to RTW and creating the action plan was simplified, and questions for assessment of goal fulfillment were added: further improvements of self-rating of health aspects were made with addition and development of items, and separate graphs for self-monitoring of health aspects and goals were developed. Continuous adjustments and refinements of the functions based on the usability test P1 were made, and further parts of SWEPPE (eg, the library, the coach function, and the user profile for the employer) were developed.

The version of SWEPPE evaluated in usability tests E1 and P2 consisted of a fully functioning web application for the employers and a smartphone app for the persons with chronic pain. The employers perceived receiving information about the employee's goal related to RTW, barriers for RTW, support wanted from the employer, and the graph to follow the employee's progress as the most useful functions in SWEPPE (Table 6). The persons with chronic pain participating in test P2 rated self-monitoring, setting a goal related to RTW, and identifying barriers and strategies as the most useful functions. Overall, the participants found SWEPPE to be helpful. For example, they thought the application was safe, provided relevant information, and would be good for many people with chronic pain. Regarding usability, the median SUS scores of the employers and persons with chronic pain were high (median 88 and 86.5, respectively; Table 6). SWEPPE was deemed to be logical and easy to use with an appealing interface. The participants in tests E1 and P2 also provided several comments regarding the different functions in SWEPPE, and the employers provided suggestions for ways to improve the application (Table 7).

Table 7. Overview of SWEPPE[a] and the modules and their content evaluated in the usability tests with the employers (test E1, n=3) and persons with chronic pain (test P2, n=5).[b]

Module in SWEPPE	Description of content	Comments from participants in tests E1 and P2
The action plan[c]	Goal setting regarding work; identification of barriers to RTW, strategies to handle these barriers, and support wanted from the employer; and weekly evaluation of work ability and fulfillment of the goal	• "Good help to set a goal with the suggestions and getting a summary in the overview" [P] • "Having a goal makes it easier to do the little extra to fulfil your wishes" [P] • "By identifying the barriers, it is easier for you to find strategies to work around them. Otherwise, it is easy to end up with bad habits and you don't know why" [P] • "I like the suggestions for strategies because many might not even think about it" [P]
Daily self-rating[c]	Self-rating of health and psychosocial aspects, work situation, and strategies	• "A very good part" [P] • "Good to be able to choose what health aspects to monitor" [P]
Self-monitoring graphs[c]	Graphs for self-monitoring health aspects, work ability, and progress toward the goal over time	• "Good with the summary in a graph" [P] • "It is easier to capture trends like not doing exercise when you have a lot of other things to do. Then you get the information in black and white that you have skipped exercise too many days and you can follow the pain curve which due to lack of exercise is getting worse" [P] • "To follow pain, stress and physical activity would help me a lot. It can help to do more exercise and it gives you a great summary if the activity helps for the pain" [P]
The coach[c]	Opportunity to ask a question and receive a written answer from a coach	• "Superb function to be able to get help via the app" [P] • "Surely good if you need support in some way like how to handle your employer" [P]
The library[c,d]	Knowledge database developed based on previous research with information (texts, films, and audio clips) that reflects a biopsychosocial perspective regarding chronic pain, physical activity, managing the situation, activity pacing, balance in daily life, sleep, workplace adaptations, tools for dialogue, and answers from the coach on common questions	• "Good texts and films. If only the employer has the time and will to learn there is a lot of good material in the app. Not only for the employer but also for me" [P] • "This would have been useful for me earlier [in the RTW process]" [P] • "I liked the library. A lot of good information" [P] • "Gathered information is always good" [E]
Shared information with the employer[c,d]	The person with chronic pain can give the employer access to the library and share information from the action plan and the graph for monitoring work ability and goal fulfillment in SWEPPE, and the employer receives the information from the parts of the action plan the employee has chosen to share; if the employee does not want to share any information from the action plan, the employer still has access to the library	• "Good and perspicuous arrangement of goal, barriers, strategies and wanted support" [E] • "Can meetings be visualized? Reconciliation meetings with the occupational health care services is an important basis that would be good to see in the graph" [E] • "It would be valuable to follow up strategies from the employee and employer that have not given results, that is changes in strategies and support wanted from the employer during rehabilitation. What has given results in the right direction and what has not" [E] • "Clearer start and goal of the weekly evaluations, it would add value if you could register concrete actions to follow up" [E] • "A simple platform for quickly finding gathered information and the employee's progress" [E] • "This is not applicable for me right now but if I would increase my working time, it would be very good to involve the employer. I think SWEPPE would be good both for me and for my employer as long as the employer has the will. The formulation in the app is clear and I think it would make communication between the employer and the employee easier" [P] • "It can be difficult to get you employer involved but with SWEPPE it can be easier for the employer to see if there is a negative trend. Unfortunately, I don't think everybody would dare to share with their employer and some employers will probably not be so engaged or even look in SWEPPE" [P] • "It's good to be able to give you employer insights about how you feel and you choose how much you want to share" [P]

[a]SWEPPE: Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers.

[b]These modules also constituted the final version of SWEPPE.

[c]Accessed by persons with chronic pain via the smartphone app.

[d]Accessed by the employers via the web application.

The final version of SWEPPE consisted of all the modules presented in Table 7. In the final version, the content of the action plan (the goal and identifying strategies and support wanted from the employer) can be individualized by the employee. The user is presented with different options (strategies or needs) to choose from, but these can be modified, and the user can also create their own options in the app. For daily self-rating of health aspects, the user is given the possibility to self-monitor not only bio-related aspects such as fatigue and pain but also psychosocial aspects such as work situation and activity balance. For daily self-rating of, for example, pain, a slider for a visual analog scale was used, ranging from 0 (no pain) to 10 (worst imaginable pain). The value is not indicated on the screen when the user is performing daily self-rating but is presented in the self-monitoring graph. Screenshots from SWEPPE are presented in Figure 3.

Figure 3. Screenshots from the final version of SWEPPE. SWEPPE is available in Swedish. Top row from left to right: the action plan (Styrkort), daily self-registration overview (Skattning) and rating of pain (Smärta), one of the library sections (Hantera din situation). Bottom row: self-monitoring (Översikt) of health aspects, work ability and goal fulfilment, employer's view of shared information from the employee.

Discussion

Principal Findings

This study describes the development of SWEPPE, a digital support application for sustainable RTW for persons with chronic pain and their employers. SWEPPE was developed with a UCD agile approach [30], where the foundation was evidence-based knowledge of chronic pain reflecting a biopsychosocial perspective and RTW in combination with involvement of the end users during the development process. To our knowledge, SWEPPE is the first eHealth intervention for both patients with chronic pain and their employers in supporting a sustainable RTW.

SWEPPE was developed by a multidisciplinary project team using a combination of the 5 principles of UCD agile design [30]. First, the separate product discovery and product creation phases provided relevant content and functions initially discovered through a workshop and a questionnaire, which led to the development of a low-fidelity prototype [39] that was tested and constituted the foundation for the product creation phase. Second, the iterative design and development used feedback from the reference groups at different stages to revise the design. Third, the software team used parallel design and development activities using one sprint ahead with scheduling and organization of the development in 2-week sprints. Fourth, continuous involvement of the end users throughout the process was ensured by using 2 reference groups who participated in questionnaires and usability testing and a user representative as part of the project team, which all provided valuable information. Fifth, artifact-mediated communication was used for the user personas and scenarios in the workshop and for the employers in the usability test. Applying these principles in combination with the competencies of the multidisciplinary project team ensured performance of systematic evaluation and development of a product relevant for the users' context. In this study, participants with chronic pain as well as the user representative emphasized the importance of presenting information in an easy, understandable way that did not require a large cognitive load. This finding is consistent with findings by Ledel Solem et al [15], where participating patients preferred a simpler presentation of content rather than gamifying design elements, as these could be challenging to use when experiencing chronic pain. The results from this study show that SWEPPE was deemed easy to use, which has been identified as a facilitator for using eHealth applications by persons with chronic pain [17].

Supporting Self-management

This study shows that SWEPPE has the potential to be a valuable tool for supporting the individual in self-management of chronic pain during the RTW process. Web-based applications or smartphone apps can be easily accessed and enable persons with chronic pain manage their condition [49] and reduce pain interference [50]. As self-management and empowerment have been identified as important parts of successful eHealth interventions [51], SWEPPE was developed to target the lack of support experienced by persons with chronic pain after finishing a rehabilitation program and where the RTW process continues [7]. Self-monitoring of health aspects was an important part to include in SWEPPE, as it is a common strategy for self-management among persons with chronic pain [17,20]. The daily self-rating in SWEPPE generated data presented in graphs for self-monitoring, a function the participants with chronic pain perceived as useful. Individuals are different in their tolerance of pain and the daily self-rating of pain in SWEPPE reflects the individual's subjective experience. The user can also choose which and how many of the biopsychosocial aspects available in SWEPPE to monitor, based on the relevance for the individual's specific situation. Self-monitoring in SWEPPE provides the user with feedback, which can contribute to learning about health aspects in relation to actions and behaviors in daily life [52]. Patients' understanding of their own self-monitoring data involves perception of the information, making inferences, and using these to change their daily activities [22], which can give them a sense of control and motivation to continue using self-management strategies [23].

Pain education is also a common part of self-management and can be related to the neuroscience of pain, medication, stress, depression, and sleep management [19]. In SWEPPE, the library was developed to provide easily accessed information about chronic pain based on a biopsychosocial perspective. The content in the library was intended to support both the employee and the employer by contributing to an increased understanding of the need to take the whole life situation into account when planning for RTW. The library was especially important for employers who wanted knowledge about how chronic pain might affect work ability and how they can support the employee during RTW [7]. Providing information through computer applications and smartphone apps has been shown to improve the level of knowledge, and the effectiveness can increase by 78% when also using at least one push notification a week [53]. In SWEPPE, a randomly selected text from the library was suggested once a week for the user in the smartphone app to inspire continuous reading. Although the library was perceived as useful by both reference groups, it was not rated as useful as other functions in SWEPPE. According to Timmers et al [53], the timing of information is crucial, as patients need to receive the right information at the right time. In this study, persons with chronic pain found that the information in the library would have been useful for them earlier in their RTW process, indicating the potential and need of SWEPPE in a clinical setting, when the users are starting their RTW process.

Regarding the dynamics of the employee-employer relationship, the participants in both reference groups were in general positive to the function of sharing information in SWEPPE to facilitate collaboration and communication between the employee and the employer. SWEPPE was built to be al tool for providing the employer with information but without the employee having to educate the employer regarding chronic pain and its consequences for work. Instead, by using SWEPPE, the employee can invite the employer and give access to the library and decide what information to share from their action plan. If the employer is willing to engage in the process and use the provided information, this could increase the employer's ability to support the employee. The issue of employees not wanting

or daring to share information with their employer was raised by some of the persons with chronic pain during usability testing. An important feature of SWEPPE was to make the employee in charge of what information to share, when and with whom to avoid the employer from using the app for control or pressure. In this study, usability tests P2 and E1 revealed the persons with chronic pain perceived the function sharing information as slightly less useful compared with the employers and compared with other functions in SWEPPE. This was due to the participating persons with chronic pain having come too far in their RTW process and commented on this function as being more useful earlier in the process. This stresses the importance of getting the user's context correct in the development of digital solutions [34]. Therefore, the usefulness of sharing information also needs to be tested further in a clinical setting to study the interplay between the employer and the employee.

Strength, Limitations, and Future Directions

There are some limitations of this study that need to be considered. SWEPPE was developed in Sweden, where the employers are prescribed by law to take actions to adapt the workplace to the individual's capacity and thereby enable the employee to RTW or stay at work. However, the rights and responsibilities of employees and employers vary among countries, and the usefulness of an app such as SWEPPE may depend on the societal system.

The number of persons in the reference group of persons with chronic pain was small and might not be representative of the whole population of persons with chronic pain. However, the participants are representative of the targeted users of SWEPPE (ie, persons with chronic pain who have participated in an IPRP and who have experiences with the RTW process). The persons with chronic pain participating in this study had come further in their RTW process than the intended users of SWEPPE. This was regarded as a strength of the study, as the participants had the experience and possibility to reflect on their needs during the RTW process and could acknowledge that SWEPPE had been useful for them earlier in the RTW process. As a result, there were lower ratings of some of the functions (eg, coach and sharing information), as these were not needed in the participants' present situation. There were also few employers participating in this study, and recruitment of employers to the reference group was more difficult, as they were experiencing a lot of time pressure. However, having employers involved in the development of SWEPPE was crucial and is a strength of this study, as they play an important role in the RTW process [54]. Overall, the small number of participants in the 2 reference groups contributed with a variety of valuable feedback relevant to the end users.

Another strength of this study was having a user representative as part of the multidisciplinary project team, as the experience-based knowledge provided by a research partner complemented the professional knowledge [55]. The user representative gave valuable feedback during the whole process on ideas, functions, and texts and helped prioritize the suggestions from the reference groups, which validated decisions made during the design and development of SWEPPE.

A strength of this study is also the use of both qualitative and quantitative methods for evaluation and feedback during the development process [56], which gave valuable information for the development of SWEPPE. During the process of creating SWEPPE, it was decided that content, functionality, and design were the most important parts to examine in usability testing and the feedback from the reference groups. Thus, not all the written texts were evaluated in the tests with the reference groups.

Usability testing requires advanced planning and involves several decisions such as selecting the setting, the tasks the user should perform, and the type of data to be collected. In this study, 3 persons with chronic pain participated in the initial usability test (P1) of the low-fidelity prototype. This may have been too few to identify all possible usability issues. It has been suggested that 5 participants are sufficient for usability testing and finding 80% of the usability problems [39]. Still, valuable information was collected during the test that confirmed that the basic structure and content in SWEPPE were in line with the users' desires and needs. The COVID-19 pandemic also influenced the options regarding testing. For example, it was not possible to conduct the tests during a physical meeting at the university. Doing a usability test on the web via a Zoom meeting might have affected the willingness for some participants to participate in the initial test. People willing to participate in a Zoom meeting might also indicate a selection bias, as these people probably were more comfortable with using technology than people who chose not to participate perhaps because they were intimidated by technology. More participants participated in the usability tests performed on the functioning smartphone apps or web applications tested at home (E1 and P2), which can be a result of the participants feeling more comfortable using their own smartphone or computer in a familiar environment [57]. These tests were performed to validate the nearly finished version of SWEPPE and to collect suggestions for further improvements. A strength of these tests was that none of the participants needed help getting started with SWEPPE.

The results of the development of SWEPPE are positive and highly usable because of the UCD agile approach. However, to investigate its effectiveness, SWEPPE needs to be tested in a clinical setting, initially in a pilot study and then in a randomized clinical trial.

Conclusions

This study reports the development of a digital support application for persons with chronic pain and their employers. SWEPPE fulfilled the need of support after IPRP with useful functions such as setting a goal related to RTW, identifying barriers and strategies for RTW, self-monitoring, and sharing information between employee and employer. The UCD agile design approach contributed to creating SWEPPE as a relevant and easy-to-use eHealth intervention. Further studies are needed to examine the effectiveness of SWEPPE in a clinical setting.

Acknowledgments

The authors would like to thank Eva Petersson, research partner from the Swedish Rheumatism Association, for her important contribution with the user's perspective throughout the process and Linköping University IT Division for valuable support, creativity, and innovative solutions during the development process. The authors would also like to thank the clinical occupational therapists at region Östergötland, Sweden, who participated in dialogues during the development of SWEPPE (Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers) to confirm the clinical relevance. This project was funded by the Swedish Research Council for Health, Working Life and Welfare (Dnr 2019-01264).

Conflicts of Interest

None declared.

References

1. Pérez C, Margarit C, Sánchez-Magro I, de Antonio A, Villoria J. Chronic pain features relate to quality of life more than physiopathology: a cross-sectional evaluation in pain clinics. Pain Pract 2017;17(7):866-878. [doi: 10.1111/papr.12533] [Medline: 27782366]
2. Gerdle B, Björk J, Henriksson C, Bengtsson A. Prevalence of current and chronic pain and their influences upon work and healthcare-seeking: a population study. J Rheumatol 2004;31(7):1399-1406. [Medline: 15229963]
3. Harker J, Reid KJ, Bekkering GE, Kellen E, Bala MM, Riemsma R, et al. Epidemiology of chronic pain in Denmark and Sweden. Pain Res Treat 2012;2012:371248 [FREE Full text] [doi: 10.1155/2012/371248] [Medline: 22693667]
4. Reid KJ, Harker J, Bala MM, Truyers C, Kellen E, Bekkering GE, et al. Epidemiology of chronic non-cancer pain in Europe: narrative review of prevalence, pain treatments and pain impact. Curr Med Res Opin 2011;27(2):449-462. [doi: 10.1185/03007995.2010.545813] [Medline: 21194394]
5. Cáceres-Matos R, Gil-García E, Barrientos-Trigo S, Porcel-Gálvez AM, Cabrera-León A. Consequences of chronic non-cancer pain in adulthood. Scoping review. Rev Saude Publica 2020;54:39 [FREE Full text] [doi: 10.11606/s1518-8787.2020054001675] [Medline: 32321056]
6. Asmundson GJ, Katz J. Understanding the co-occurrence of anxiety disorders and chronic pain: state-of-the-art. Depress Anxiety 2009;26(10):888-901. [doi: 10.1002/da.20600] [Medline: 19691031]
7. Svanholm F, Liedberg GM, Löfgren M, Björk M. Factors of importance for return to work, experienced by patients with chronic pain that have completed a multimodal rehabilitation program - a focus group study. Disabil Rehabil 2020:1-9. [doi: 10.1080/09638288.2020.1780479] [Medline: 32580604]
8. Magalhães L, Chan C, Chapman AC, Majed L, Samigullina R, Trninic D, et al. Successful return to work of individuals with chronic pain according to health care providers: a meta-synthesis. Cad Bras Ter Ocup 2017;25(4):825-837. [doi: 10.4322/2526-8910.ctoar2063]
9. Cohen D, Allen J, Rhydderch M, Aylward M. The return to work discussion: a qualitative study of the line manager conversation about return to work and the development of an educational programme. J Rehabil Med 2012;44(8):677-683 [FREE Full text] [doi: 10.2340/16501977-0996] [Medline: 22729796]
10. Saunders SL, Nedelec B. What work means to people with work disability: a scoping review. J Occup Rehabil 2014;24(1):100-110. [doi: 10.1007/s10926-013-9436-y] [Medline: 23519737]
11. Grant M, O-Beirne-Elliman J, Froud R, Underwood M, Seers K. The work of return to work. Challenges of returning to work when you have chronic pain: a meta-ethnography. BMJ Open 2019;9(6):e025743 [FREE Full text] [doi: 10.1136/bmjopen-2018-025743] [Medline: 31227529]
12. Scott W, Chilcot J, Guildford B, Daly-Eichenhardt A, McCracken LM. Feasibility randomized-controlled trial of online Acceptance and Commitment Therapy for patients with complex chronic pain in the United Kingdom. Eur J Pain 2018;22(8):1473-1484. [doi: 10.1002/ejp.1236] [Medline: 29704880]
13. Coutu MF, Légaré F, Durand MJ, Stacey D, Labrecque ME, Corbière M, et al. Acceptability and feasibility of a shared decision-making model in work rehabilitation: a mixed-methods study of stakeholders' perspectives. J Occup Rehabil 2019;29(1):128-139. [doi: 10.1007/s10926-018-9770-1] [Medline: 29663111]
14. Braillard O, Cedraschi C, Jesaimani A, Piguet V. Chronic noncancer pain and patient education: a place for e-learning? Rev Med Suisse 2015;11(480):1400-1405. [Medline: 26267947]
15. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Børøsund E, Schreurs KM, et al. A user-centered approach to an evidence-based electronic health pain management intervention for people with chronic pain: design and development of EPIO. J Med Internet Res 2020;22(1):e15889 [FREE Full text] [doi: 10.2196/15889] [Medline: 31961331]
16. Kristjánsdóttir Ó, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain. Part 2: 11-month follow-up results of a randomized trial. J Med Internet Res 2013;15(3):e72 [FREE Full text] [doi: 10.2196/jmir.2442] [Medline: 23538392]

17. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Mirkovic J, Børøsund E, et al. Patients' needs and requirements for eHealth pain management interventions: qualitative study. J Med Internet Res 2019;21(4):e13205 [FREE Full text] [doi: 10.2196/13205] [Medline: 30877780]
18. Hadjistavropoulos HD, Faller YN, Klatt A, Nugent MN, Dear BF, Titov N. Patient perspectives on strengths and challenges of therapist-assisted internet-delivered cognitive behaviour therapy: using the patient voice to improve care. Community Ment Health J 2018;54(7):944-950 [FREE Full text] [doi: 10.1007/s10597-018-0286-0] [Medline: 29804226]
19. Devan H, Farmery D, Peebles L, Grainger R. Evaluation of self-management support functions in apps for people with persistent pain: systematic review. JMIR Mhealth Uhealth 2019;7(2):e13080 [FREE Full text] [doi: 10.2196/13080] [Medline: 30747715]
20. Devan H, Hale L, Hempel D, Saipe B, Perry MA. What works and does not work in a self-management intervention for people with chronic pain? Qualitative systematic review and meta-synthesis. Phys Ther 2018;98(5):381-397. [doi: 10.1093/ptj/pzy029] [Medline: 29669089]
21. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
22. Mamykina L, Smaldone AM, Bakken SR. Adopting the sensemaking perspective for chronic disease self-management. J Biomed Inform 2015;56:406-417 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.006] [Medline: 26071681]
23. Morton K, Dennison L, May C, Murray E, Little P, McManus RJ, et al. Using digital interventions for self-management of chronic physical health conditions: a meta-ethnography review of published studies. Patient Educ Couns 2017;100(4):616-635 [FREE Full text] [doi: 10.1016/j.pec.2016.10.019] [Medline: 28029572]
24. Thurnheer SE, Gravestock I, Pichierri G, Steurer J, Burgstaller JM. Benefits of mobile apps in pain management: systematic review. JMIR Mhealth Uhealth 2018;6(10):e11231 [FREE Full text] [doi: 10.2196/11231] [Medline: 30348633]
25. Pfeifer AC, Uddin R, Schröder-Pfeifer P, Holl F, Swoboda W, Schiltenwolf M. Mobile application-based interventions for chronic pain patients: a systematic review and meta-analysis of effectiveness. J Clin Med 2020;9(11):3557 [FREE Full text] [doi: 10.3390/jcm9113557] [Medline: 33167300]
26. Du S, Liu W, Cai S, Hu Y, Dong J. The efficacy of e-health in the self-management of chronic low back pain: a meta analysis. Int J Nurs Stud 2020;106:103507. [doi: 10.1016/j.ijnurstu.2019.103507] [Medline: 32320936]
27. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: a systematic review. Best Pract Res Clin Rheumatol 2016;30(6):1098-1109. [doi: 10.1016/j.berh.2017.04.002] [Medline: 29103552]
28. Coe-O'Brien R, Joseph L, Kuisma R, Paungmali A, Sitilertpisan P, Pirunsan U. Outcome measures used in the smartphone applications for the management of low back pain: a systematic scoping review. Health Inf Sci Syst 2020;8(1):5 [FREE Full text] [doi: 10.1007/s13755-019-0097-x] [Medline: 31938540]
29. Ghazali M, Ariffin NA, Omar R. User centered design practices in healthcare: a systematic review. In: 3rd International Conference on User Science and Engineering. 2014 Presented at: i-USEr '14; September 2-5, 2014; Shah Alam, Malaysia p. 91-96. [doi: 10.1109/iuser.2014.7002683]
30. Brhel M, Meth H, Maedche A, Werder K. Exploring principles of user-centered agile software development: a literature review. Inf Softw Technol 2015;61(C):163-181. [doi: 10.1016/j.infsof.2015.01.004]
31. Jurca G, Hellmann TD, Maurer F. Integrating agile and user-centered design: a systematic mapping and review of evaluation and validation studies of agile-UX. In: Proceedings of the 2014 Agile Conference. 2014 Presented at: AGILE '14; July 28-August 1, 2014; Kissimmee, FL p. 24-32 URL: https://www.agilealliance.org/wp-content/uploads/2016/08/Researh.2014.Jurca_.Hellmann.Maurer.pdf [doi: 10.1109/agile.2014.17]
32. da Silva TS, Marin A, Maurer F, Silveira M. User-centered design and agile methods: a systematic review. In: 2011 Agile Conference. 2011 Presented at: AGILE '11; August 7-13, 2011; Salt Lake City, UT p. 77-86. [doi: 10.1109/agile.2011.24]
33. Abelein U, Paech B. Understanding the influence of user participation and involvement on system success – a systematic mapping study. Empir Software Eng 2015;20(1):28-81. [doi: 10.1007/s10664-013-9278-4]
34. Buis LR, Huh-Yoo J. Common shortcomings in applying user-centered design for digital health. IEEE Pervasive Comput 2020;19(3):45-49. [doi: 10.1109/mprv.2020.2997615]
35. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
36. Wynne-Jones G, Artus M, Bishop A, Lawton SA, Lewis M, Jowett S, SWAP Study Team. Effectiveness and costs of a vocational advice service to improve work outcomes in patients with musculoskeletal pain in primary care: a cluster randomised trial (SWAP trial ISRCTN 52269669). Pain 2018;159(1):128-138. [doi: 10.1097/j.pain.0000000000001075] [Medline: 28976423]
37. Krippendorff K. Content analysis: an introduction to its methodology. 3 edition. Thousand Oaks, CA: Sage Publications; 2013.
38. Kiernan L, McMahon M. How scenarios support empathy in design: a case study of undergraduate students. Iterations 2015;1:10-13. [doi: 10.31880/10344/5861]

39. Lazar J, Feng JH, Hochheiser H. Usability testing. In: Lazar J, Feng JH, Hochheiser H, editors. Research methods in human-computer interaction. 2 edition. Cambridge, MA: Elsevier; 2017:263-298.
40. Joe J, Chaudhuri S, Le T, Thompson H, Demiris G. The use of think-aloud and instant data analysis in evaluation research: exemplar and lessons learned. J Biomed Inform 2015;56:284-291 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.001] [Medline: 26071683]
41. Brown JL. Empathy mapping: a guide to getting inside a user's head. UX Booth. 2018. URL: https://www.uxbooth.com/articles/empathy-mapping-a-guide-to-getting-inside-a-users-head/ [accessed 2021-02-23]
42. Sowden G, Main CJ, van der Windt DA, Burton K, Wynne-Jones G. The development and content of the vocational advice intervention and training package for the study of work and pain (SWAP) trial (ISRCTN 52269669). J Occup Rehabil 2019;29(2):395-405 [FREE Full text] [doi: 10.1007/s10926-018-9799-1] [Medline: 29982957]
43. Nieuwenhuizen MG, de Groot S, Janssen TW, van der Maas LC, Beckerman H. Canadian occupational performance measure performance scale: validity and responsiveness in chronic pain. J Rehabil Res Dev 2014;51(5):727-746 [FREE Full text] [doi: 10.1682/JRRD.2012.12.0221] [Medline: 25357091]
44. Wressle E, Lindstrand J, Neher M, Marcusson J, Henriksson C. The Canadian occupational performance measure as an outcome measure and team tool in a day treatment programme. Disabil Rehabil 2003;25(10):497-506. [doi: 10.1080/0963828031000090560] [Medline: 12745961]
45. Dear BF, Gandy M, Karin E, Fogliati R, Fogliati VJ, Staples LG, et al. The pain course: 12- and 24-month outcomes from a randomized controlled trial of an internet-delivered pain management program provided with different levels of clinician support. J Pain 2018;19(12):1491-1503. [doi: 10.1016/j.jpain.2018.07.005] [Medline: 30099209]
46. Brooke J. SUS - a quickdirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability evaluation in industry. London, UK: Taylor and Francis; 1996:189-194.
47. Lewis JR, Utesch BS, Maher DE. Measuring perceived usability: the SUS, UMUX-LITE, and AltUsability. Int J Hum-Comput Interact 2015;31(8):496-505. [doi: 10.1080/10447318.2015.1064654]
48. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
49. Garg S, Garg D, Turin TC, Chowdhury MF. Web-based interventions for chronic back pain: a systematic review. J Med Internet Res 2016;18(7):e139 [FREE Full text] [doi: 10.2196/jmir.4932] [Medline: 27460413]
50. Slattery BW, Haugh S, O'Connor L, Francis K, Dwyer CP, O'Higgins S, et al. An evaluation of the effectiveness of the modalities used to deliver electronic health interventions for chronic pain: systematic review with network meta-analysis. J Med Internet Res 2019;21(7):e11086 [FREE Full text] [doi: 10.2196/11086] [Medline: 31317869]
51. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
52. Ruckenstein M. Visualized and interacted life: personal analytics and engagements with data doubles. Societies 2014;4(1):68-84. [doi: 10.3390/soc4010068]
53. Timmers T, Janssen L, Kool RB, Kremer JA. Educating patients by providing timely information using smartphone and tablet apps: systematic review. J Med Internet Res 2020;22(4):e17342 [FREE Full text] [doi: 10.2196/17342] [Medline: 32281936]
54. Liedberg GM, Björk M, Dragioti E, Turesson C. Qualitative evidence from studies of interventions aimed at return to work and staying at work for persons with chronic musculoskeletal pain. J Clin Med 2021;10(6):1247 [FREE Full text] [doi: 10.3390/jcm10061247] [Medline: 33802906]
55. Schipper K, Abma TA, van Zadelhoff E, van de Griendt J, Nierse C, Widdershoven GA. What does it mean to be a patient research partner? An ethnodrama. Qual Inq 2010;16(6):501-510. [doi: 10.1177/1077800410364351]
56. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
57. Hass C. A practical guide to usability testing. In: Edmunds M, Hass C, Holve E, editors. Consumer informatics and digital health: solutions for health and health care. Berlin, Germany: Springer; 2019:107-124.

Abbreviations

IPRP: interdisciplinary pain rehabilitation program
RTW: return to work
SUS: System Usability Scale
SWEPPE: Sustainable Worker Digital Support for Persons With Chronic Pain and Their Employers
UCD: user-centered design

Edited by A Kushniruk; submitted 14.09.21; peer-reviewed by N Khalili-Mahani, D Garg; comments to author 28.11.21; accepted 07.12.21; published 14.03.22.

Please cite as:
Turesson C, Liedberg G, Björk M
Development of a Digital Support Application With Evidence-Based Content for Sustainable Return to Work for Persons With Chronic Pain and Their Employers: User-Centered Agile Design Approach
JMIR Hum Factors 2022;9(1):e33571
URL: https://humanfactors.jmir.org/2022/1/e33571
doi:10.2196/33571
PMID:35285814

Original Paper

User Engagement and Assessment of Treatment Effectiveness in Patients Using a Novel Digital mHealth App During Spinal Cord Stimulation Screening Trials

Jennifer M Lee[1], MD; Rex Woon[2], MSc; Mandy Ramsum[2], BSc; Daniel S Halperin[2], PhD; Roshini Jain[2], MSc

[1]EvergreenHealth Pain Care, Kirkland, WA, United States

[2]Boston Scientific Neuromodulation, Valencia, CA, United States

Corresponding Author:
Roshini Jain, MSc
Boston Scientific Neuromodulation
25155 Rye Canyon Loop
Valencia, CA, 91355
United States
Phone: 1 66169494350
Email: Roshini.Jain@bsci.com

Abstract

Background: Patient outcomes and experience during a Spinal Cord Stimulation (SCS) screening trial can have a significant effect on whether to proceed with long-term, permanent implantation of an SCS device for the treatment of chronic pain. Enhancing the ability to track and assess patients during this initial trial evaluation offers the potential for improved understanding regarding the suitability of permanent device implantation as well as identification of the SCS-based neurostimulative modalities and parameters that may provide substantial analgesia in a patient-specific manner.

Objective: In this report, we aimed to describe a preliminary, real-world assessment of a new, real time tracking, smart, device-based digital app used by patients with chronic pain undergoing trial screening for SCS therapy.

Methods: This is a real-world, retrospective evaluation of 13,331 patients diagnosed with chronic pain who used the new "mySCS" mobile app during an SCS screening trial. The app design is health insurance portability and accountability act (HIPAA)-compliant and compatible with most commercially available smartphones (eg, Apple, iPhone, and Android). The app enables tracking of user-inputted health-related responses (ie, pain relief, activity level, and sleep quality) in addition to personal trial goals and a summary of overall experience during the SCS trial. A deidentified, aggregate analysis of user engagement, user-submitted responses, and overall trial success was conducted.

Results: When provided the opportunity, the percentage of users who engaged with the tracking app for ≥50% of the time during their trial was found to be 64.43% (n=8589). Among the 13,331 patients who used the app, 58.24% (n=7764) entered a trial goal. Most patients underwent SCS screening with a trial duration of at least 7 days (n=7739, 58.05%). Of those patients who undertook a 7-day SCS trial, 62.30% (n=3456) engaged the app for 4 days or more. In addition, among all who submitted descriptive responses using the app, health-related improvements were reported by 77.84% (n=10,377) of patients who reached day 3 of the screening phase assessment and by 83.04% (n=11,070) of those who reached trial completion. A trial success rate of 91% was determined for those who used the app (versus 85% success rate for nonusers).

Conclusions: Data from this initial, real-world examination of a mobile, digital-health–based tracking app ("mySCS"), as used during the SCS screening phase, demonstrate that substantial patient engagement can be achieved while also providing for the acquisition of more real time patient-outcome measures that may help facilitate improved SCS trial success.

(JMIR eHum Factors 2022;9(1):e35134) doi:10.2196/35134

KEYWORDS

spinal cord stimulation; SCS; chronic pain; digital health; smartphone app; mobile health; mHealth; smart device; digital application; application; app; spine

Introduction

Widespread use of smart devices (ie, mobile phones and tablets) has fostered an unprecedented growth in the use of digital-based platforms and apps enabling real time tracking of health-related outcomes and experiences of patients undergoing treatment of chronic pain [1-3]. More importantly, these contemporary tools have been demonstrated to help instill a greater level of assurance in patients, that their pain condition can be controlled. These devices can also promote improved self-monitoring of the multidimensional experience of chronic pain, and some concurrently allow health care providers enhanced connectivity to real time outcomes reported by those under their care [4-6]. The implementation and use of new digital methodologies in the context of health care is underpinned by what is now referred to as ecological momentary assessment (EMA), an approach constituted by the frequent sampling of study subjects' behavior, outcomes, and experiences in real time within the real-world environments in which they live and participate [7,8]. EMA, as a technique that relies on the repeated collection of data pertaining to the health-related condition of the patient, is therefore thought to reduce recall bias and enable improved assessment of the experience of patients with chronic pain. As such, patient-specific information, acquired in a spatiotemporal manner using EMA-based methods, may help provide better clinical assessment of individual patients given the highly subjective and variable experience of those having various chronic pain disorders.

Spinal Cord Stimulation (SCS) is an established treatment option for chronic pain, which delivers electrical impulses to neuronal tissues within or adjacent to the spinal cord. These impulses are capable of interrupting the transmission of dysregulated pain signals, typically due to nerve injury, which can occur between one or more localized anatomical areas (eg, low back, leg, foot, upper limb, etc) and the human brain. Typically, before patients are permanently implanted with an SCS device, they must first undergo a screening period, commonly termed as a trial, where they experience SCS therapy for a short duration of time (eg, up to 3-7 days) to assess whether the applied treatment is effective for reducing pain. Only on the basis of the experience and success of the trial, as determined by verbal reporting of significant pain relief (defined as a 50% or greater reduction in pain) as well as improvement in function, is a permanent device implanted for long-term use. However, SCS screening trials can be challenging given the difficulty that some patients have maintaining the engagement necessary to sufficiently assess whether their pain relief and functional goals are being effectively met or not due to complications or lack of successful outcomes, which are known to be associated with higher costs resulting from repeated attempts at management of chronic pain [9-12]. Thus, whether the ability to record and assess patient feedback in real time during this initial screening phase could allow for an improved experience for patients is an open question. As noted, various publications have previously reviewed and examined the use of mobile, digital health-based apps in patients with chronic pain [13]. However, to our knowledge, no published reports describing the use of a mobile, digital companion app during the SCS trial phase exist in the peer-reviewed literature. Here, we describe our initial, real-world evaluation of a simple, mobile, smart device-based app implemented as a tool to track user-submitted goals, health-related assessments, and satisfaction in those undergoing SCS screening in order to provide a more real time examination of the trial experience for patients with chronic pain. In so doing, we sought to also determine how capable the app is in eliciting patient engagement during this phase and whether this metric could be potentially correlative with SCS trial success.

Methods

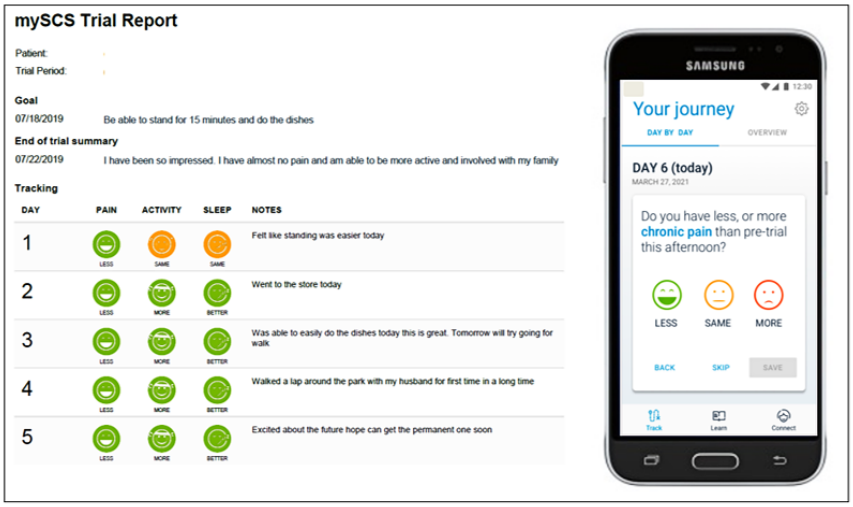
The newly commercially available mySCS app (Boston Scientific) was provided at no cost to patients diagnosed with chronic pain who participated in a trial (with a duration length up to 10 days) of an SCS system. The patients were invited to use the app either before or at their trial appointment. The patients were informed of the app either from a product brochure, their physician or physician office, or a company representative (Boston Scientific). The patients were presented with the opportunity to download the app during their SCS trial, but it was not a required condition in order to undergo their SCS trial. The patients were directed to carry out one of the following to help with the downloading process: search the app store to download, use a provided QR code to scan, or use an activation link. The QR code and activation links would take patients directly to the app listing on the app store to download the app. Company representatives were available at the trial appointment to assist with downloading if needed, but most patients were instructed to complete the installation before the day of their trial. The app was designed to be health insurance portability and accountability act (HIPAA)-compliant and was installed onto each participating patient's personal smartphone (eg, Apple iPhone and Android) and is compatible with most recent smartphones and tablets. However, Android 8.0 or above and iOS 11 or later for iPhone, iPad, and iPod touch are required. The app is currently available on smart tablets, but not smart watches. All patients were required to provide consent to the terms of use following download and set up of the app. Figure 1 provides a pictorial representation of the app interface and a sample trial report that can be generated daily or at the end of the trial. In order to be eligible for inclusion, all patients were required to be least 18 years of age or older and have the following baseline demographic information available: trial start and end date, age, gender, and trial status (listed either as a "successful" or "failed" trial). Those patients listed with an "inconclusive" trial status and patients who underwent multiple trials were excluded from data analysis. Patients who were provided the app were instructed to use the app daily to record progress and their personal experience during the SCS trial. Patient-entered information was stored to a secure database that allows for exporting in PDF following completion of the trial. SCS screening trial data from a separate cohort of patients who did not use the app was also collected for comparative assessment. Gender and age demographic information was collected.

The mySCS app enables convenient tracking of information entered directly into the app by the patient. The patients were prompted to enter an assessment each day of their trial, which

was of variable length, as determined by their physician. Typical trials range from 3 to 7 days, and the app reminds the patient to enter an assessment each day. Categorical descriptors selected in real time by patients were used to track ongoing user-submitted responses on a daily basis (vs patient's pretrial condition) including intensity of chronic pain ("less," "same," and "more"), level of activity ("less," "same," and "more"), and sleep quality ("worse," "same," and "better"). Additionally, the patients entered their personal trial goal and overall trial satisfaction. User engagement with the digital app was defined as any user-submitted response, comment, goal, or assessment into the app. For each patient using the app, the number of days engaged with the app over the course of the length of the trial period was determined. Content analysis was carried out by assessing the frequency of terms entered into the app by patients. Recorded patient goals and summaries of trial satisfaction were assessed following completion of the SCS trial. Further, patient-entered trial satisfaction summaries were evaluated using a bigram analysis (occurrence of 2 consecutive words as a pair).

A successful SCS trial is conventionally defined as a ≥50% reduction in pain intensity at the end of the trial (vs pretrial pain intensity). Relative improvement in trial success between those who used the app (for at least 1 day) versus those who did not use the app was calculated by comparing the trial success rate between both of these separate cohorts using a one-tailed chi-squared test to determine if the proportions were different from each other at a statistical significance level of 0.05. All analyses were performed in Python (Python Software Foundation). The Pandas package was used for data management, and the Scipy package was used for performing statistical tests.

Figure 1. Pictorial representation of the mySCS app interface and sample trial report.



Ethics Consideration

All data were collected in aggregate (independent of the centers in which patients were implanted) and was obtained fully deidentified, thereby obviating the need for ethics board review approval of this evaluation per United States 45 Code of Federal Regulations (CFR) § 46.104(d)(2)(i).

Results

Data were acquired from a cohort of 13,331 patients who used the new mobile app during their SCS trial. Additionally, data from 12,196 patients who did not use the app were also obtained and evaluated. Gender and age demographics for those patients who did or did not use the app are shown in Table 1.

Table 1. Patient demographics (significant difference with chi-square test $P<.001$).

Characteristics	Patients with the mySCS app (n=13,331)	Patients without the mySCS app (n=12,196)
Gender—male, n (%)	5760 (43)	5648 (46)
Age (years), mean (SD)	60 (14)	66 (13)
Trial success, %	91	85

In total, 58.05% (n=7739) of those patients who had access and used the app underwent at least a 7-day trial with a maximum duration of up to 10 days. Initial engagement with the app required patients to enter in a personal goal for their SCS trial. Of the patients who used the app, nearly 58.24% (n=7764) were noted to have entered a trial goal. Analysis of the most prevalent key health-related functional terms occurring within the text of goals entered by patients using the mobile app is depicted in Table 2.

The most common term ("walk") was found in 46.81% (n=3634) of the entries provided by patients, followed by "less pain" (n=3514, 45.26%). Analysis of user engagement demonstrated that 64.43% (n=8589) of all users engaged the tracking app for at least 50% of the time within the total duration of their screening trial (Figure 2). Trials carried out for 7-days in duration were found to have been undertaken most frequently among those who used the app. Analysis of app engagement among those in this subcohort revealed that 62.3% (n=3456) of

those who successfully completed a 7-day trial engaged the app for 4 days or more (Table 3).

Among all patients who used the app through day 3 and out to trial completion, 77.84% (n=10377) and 83.04% (n=11,070) demonstrated improvement in health-related metrics (ie, pain, activity level, and sleep quality), respectively. Of those patients who did not use the app, an 85% trial success rate (ie, ≥50% pain relief) was noted. Alternatively, a trial success rate of 91% was determined among those who did use the tracking app (Table 4) representing a 6% increase in trial success.

Evaluation of trial satisfaction summaries submitted by 1535 patients who used the app was conducted using a bigram analysis of the content that was recorded into the app following completion of the trial (Table 5). The most common consecutive 2-word phrase entered into the app by patients was found to be "less pain" followed by "more active" and "very well."

Table 2. Occurrence of key terms in patient-entered goals (n=7795)

Key terms	Values, n (%)
Walk	3634 (46.81)
Less pain	3514 (45.26)
Sleep	1520 (19.58)
Stand	1265 (16.29)
Sit	730 (9)

Figure 2. App user engagement during trial period.

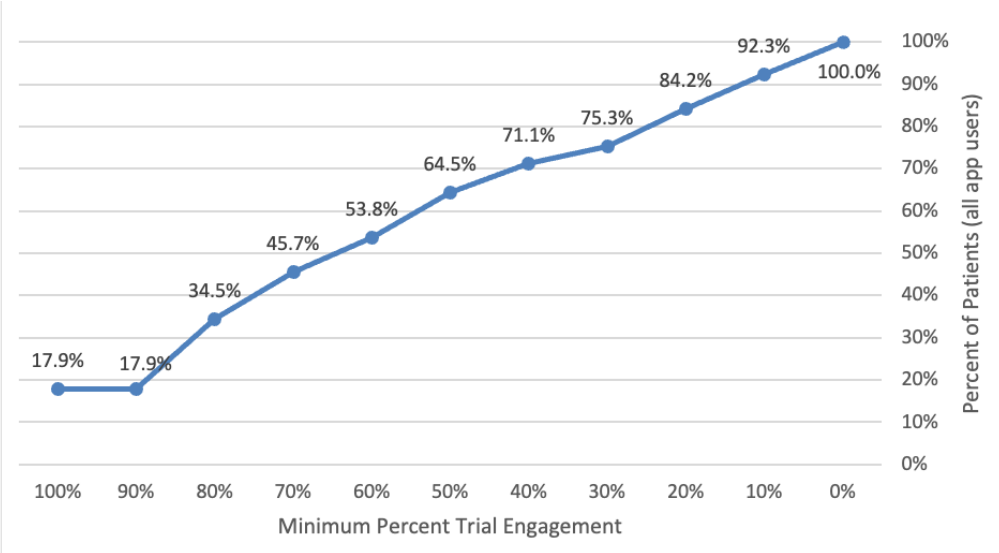


Table 3. App engagement duration among users undergoing a 7-day trial (n=5547).

App engagement duration (days)	Patients with a 7-day trial, n (%)
≤1	954 (17.20)
2-3	1143 (20.61)
≥4	3450 (62.19)

Table 4. Trial success (≥50% pain relief) among app users and nonusers.

	Patients with successful trial, n (%)
Patients with the mySCS app (n=13,331)	12,133 (91.01)
Patients without the mySCS app (n=12,196)	10,359 (84.94)

Table 5. Occurrence of bi-grams in patient-entered trial assessments (n=1535).

Bi-gram	Trial assessments, n (%)
Less pain	108 (7.04)
More active	95 (6.19)
Very well	68 (4.43)
So much	62 (4.04)
Pain relief	59 (3.84)
Much better	57 (3.71)

Discussion

Principal Findings

This report, to our knowledge, is the first to demonstrate that successful patient engagement and capability for real time assessment of SCS trial outcomes can be achieved using a new, HIPAA-compliant, digital health-based mobile tracking app ("mySCS"). In addition, the use of this app was found to be associated with an increased proportion of SCS trial success. We also observed that patients using this new app on their smart device were able to better recognize their responses to treatment (compared to patient recall), while also enabling the progress of patients to be more clearly communicated after the trial period. This is a potentially important finding given the lack of reliability and the well-known difficulty that patients are known to have when using memory-based recall to estimate the intensity of the pain they experience [14-16].

In accordance with previous reports that have shown that patients with chronic pain are strongly interested and eager to use a digital health-based mobile app, we found that a large percentage of patients undergoing an SCS trial, if offered, are willing to use such a tool [4,6,8]. Given the importance the trial phase can have on the decision to proceed with permanent implantation of an SCS device, this sizable rate of user participation demonstrates the potential viability of a mobile tracking app to improve the utility of an SCS trial. This result is encouraging given that a substantial portion of the SCS-implanted population is over the age of 65 years and often beset by psychosocial factors that can affect patient function shown to be correlated with less successful trial outcomes [17,18]. Furthermore, in surveys conducted prior to the COVID-19 pandemic, those aged 65 years or older were found to have had substantially less knowledge or experience with digital health apps [19-21]. Interestingly, in accordance with more recent reports assessing user engagement with virtual health tools during the COVID-19 pandemic, 2 published studies by Han et al [22] and Lu et al [23] (both of which were conducted during the height of the recent COVID-19 pandemic) reported the successful implementation of remote monitoring and programming of those implanted with an SCS device; they also found that an overwhelming majority of patients had a strong desire for follow-up visits that do not require in-person interaction with their health care provider [22-25]. The results from these early investigations and of those now presented in this report thus jointly suggest that the use of digital health technologies that can track and record SCS outcomes and experience are likely to be highly desired and used by patients, including potentially even those >65 years of age, as a consequence, at least in part, of the necessities imposed on older adults to become more accustomed to remote-accessible tools during the COVID-19 pandemic. Moreover, it has also been postulated that costs associated with the implementation of SCS for treatment of chronic pain may be lowered when using digital health tools that allow for remote-based patient tracking and follow-up visits [22,23]. Though clinical examination (and publication of data) pertaining to the use of digital health-based tools in the context of SCS therapies is still quite limited at this time, these initial reports would suggest that the benefit to patients and providers (with the potential of integrating the use of new digital technologies as a part of an SCS-based therapeutic regimen for chronic pain) could be potentially substantial. We further surmise that the gradual shift toward more ubiquitous use of various digital tools in the real-world clinical setting may likely facilitate the eventual incorporation of digital health-based technology as a key component of the routine care provided to patients within the practical context of interventional pain management (to better monitor and treat those implanted with an SCS device in a more remote and personalized manner).

Patients using the new mobile app examined in this evaluation demonstrated an improved rate of trial success (vs a separate cohort of those who did not use the app). However, it is currently unknown if patient satisfaction at completion of an SCS trial is in fact greatly altered in patients implanted with an SCS device for chronic pain who use a digital health-based tool versus those who do not. Nonetheless, an extrapolation of the data described in Table 4 of this report indicates that, had those who did not use the app during their SCS trial chosen to do so, up to an additional 732 patients (calculated per the difference between 91% of app users and 85% of non–app users who had trial success) could have potentially achieved a successful trial. Interestingly, other reports of patients using implantable systems, such as Deep Brain Stimulation devices, have shown that a measurable improvement in outcomes (eg, patient satisfaction) can be observed in those undergoing remote monitoring using digital health technologies [26-28]. These notable improvements are thought to be due, at least in part, to an increase in the positive impression of treatment and overall psychological benefit that patients obtain when they log their progress routinely and reflect on their current health-related state. Mental health and treatment expectations are thought to have at least some effect on outcomes in most patients treated with SCS for chronic pain [18,29,30]. Therefore, whether the use of available digital health tools equipped with EMA-based tracking or remote-based

communication features is, in turn, correlated with improved clinical outcomes (eg, psychological health measures) is now an important question warranting further investigation.

Limitations

Given the preliminary nature of the evaluation described, limitations associated with the analyses described in this report must be noted. First, assessment of data was conducted retrospectively on the basis of the initial real-world launch of the tracking app made available to SCS-implanted patients. Future investigations are now needed to prospectively examine the impact of new digital health tools on patient outcomes using predefined measures and study designs including those that address the role that treatment expectations (ie, placebo responses) may have on obtained clinical outcomes. Additionally, the version of the app used by those described in this report did not allow for a highly detailed recording of baseline demographics, procedural information, or pain intensity based on an established rating scale (eg, Visual Analogue Scale). Going forward, procurement of such patient information may facilitate the detection of any selection bias (ie, bias as a result of the inadvertent selection for a particular patient segment within the overall cohort of assessed individuals such that the sample evaluated is not truly representative of the actual intended patient population) among those who used the app versus those who did not use the app. Information as it pertains to medical history, lead location, spinal level placement, and applied stimulation parameters of those who used the app may also have provided further insight as to the presence of any correlations associated with patient engagement and trial success. Moreover, no data were available that would have allowed an analysis of the percentage of patients who continued to receive a permanent implant after their successful trial or whether specific goals beyond pain relief alone (eg, improvements in functional disability) were achieved. Rates of conversion from trial to permanent implant are thought to have implications regarding overall device efficacy as well as other aspects related to the successful use of SCS as a therapeutic option for chronic pain [31]. As such, examination of this key metric in patients using digital health tools as part of an SCS trial, such as a future version of the mobile tracking app described in this report, is now warranted.

Conclusion

This initial, real-world examination of a real time, mobile, digital-health–based tracking app (“mySCS”), as used during the SCS trial, demonstrates that substantial patient engagement can be achieved while also providing for more reliable and quantitative outcome measures that may help facilitate increased SCS trial success. The use of a novel digital-health–based mobile app therefore may constitute an important new approach toward fostering an improved experience during the SCS trial. A greater understanding of patient-specific clinical responses may also allow for better decision-making and evaluation regarding the appropriate use and effectiveness of SCS as a therapeutic strategy for treatment of chronic pain. Additional study and assessment are now needed to further understand the potential benefits of digital-health–based tools in the context of SCS therapy.

Authors' Contributions

RW, MR, and RJ initially conceived of this work. All authors provided intellectual input regarding the design of data collection and analysis. RW conducted all biostatistical calculations. DSH wrote the first draft of the manuscript with assistance from RW. All authors contributed to the editing and critical revision of the manuscript. The final version of the manuscript was approved by all authors.

Conflicts of Interest

This work is sponsored by Boston Scientific. JML is a consultant for Boston Scientific. All other authors are salaried employees of Boston Scientific.

References

1. Lalloo C, Shah U, Birnie KA, Davies-Chalmers C, Rivera J, Stinson J, et al. Commercially Available Smartphone Apps to Support Postoperative Pain Self-Management: Scoping Review. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e162 [FREE Full text] [doi: 10.2196/mhealth.8230] [Medline: 29061558]
2. Slepian PM, Peng M, Janmohamed T, Kotteeswaran Y, Manoo V, Blades AM, et al. Engagement with Manage My Pain mobile health application among patients at the Transitional Pain Service. Digit Health 2020 Oct 13;6:2055207620962297 [FREE Full text] [doi: 10.1177/2055207620962297] [Medline: 33117557]
3. Thurnheer SE, Gravestock I, Pichierri G, Steurer J, Burgstaller JM. Benefits of Mobile Apps in Pain Management: Systematic Review. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e11231 [FREE Full text] [doi: 10.2196/11231] [Medline: 30348633]
4. Suso-Ribera C, Castilla D, Zaragozá I, Ribera-Canudas M, Botella C, García-Palacios A. Validity, Reliability, Feasibility, and Usefulness of Pain Monitor: A Multidimensional Smartphone App for Daily Monitoring of Adults With Heterogenous Chronic Pain. Clin J Pain 2018 Oct;34(10):900-908. [doi: 10.1097/AJP.0000000000000618] [Medline: 29659375]
5. Stinson J, Jibb L, Nguyen C, Nathan P, Maloney A, Dupuis L, et al. Construct validity and reliability of a real-time multidimensional smartphone app to assess pain in children and adolescents with cancer. Pain 2015 Dec;156(12):2607-2615. [doi: 10.1097/j.pain.0000000000000385] [Medline: 26580680]

6. Galve Villa M, S Palsson T, Cid Royo A, R Bjarkam C, Boudreau SA. Digital Pain Mapping and Tracking in Patients With Chronic Pain: Longitudinal Study. J Med Internet Res 2020 Oct 26;22(10):e21475 [FREE Full text] [doi: 10.2196/21475] [Medline: 33104012]
7. Ono M, Schneider S, Junghaenel DU, Stone AA. What Affects the Completion of Ecological Momentary Assessments in Chronic Pain Research? An Individual Patient Data Meta-Analysis. J Med Internet Res 2019 Feb 05;21(2):e11398 [FREE Full text] [doi: 10.2196/11398] [Medline: 30720437]
8. Bhatia A, Kara J, Janmohamed T, Prabhu A, Lebovic G, Katz J, et al. User Engagement and Clinical Impact of the Manage My Pain App in Patients With Chronic Pain: A Real-World, Multi-site Trial. JMIR Mhealth Uhealth 2021 Mar 04;9(3):e26528 [FREE Full text] [doi: 10.2196/26528] [Medline: 33661130]
9. Eldabe S, Duarte RV, Gulve A, Thomson S, Baranidharan G, Houten R, et al. Does a screening trial for spinal cord stimulation in patients with chronic pain of neuropathic origin have clinical utility and cost-effectiveness (TRIAL-STIM)? A randomised controlled trial. Pain 2020 Dec 29;161(12):2820-2829 [FREE Full text] [doi: 10.1097/j.pain.0000000000001977] [Medline: 32618875]
10. Chadwick R, McNaughton R, Eldabe S, Baranidharan G, Bell J, Brookes M, et al. To Trial or Not to Trial Before Spinal Cord Stimulation for Chronic Neuropathic Pain: The Patients' View From the TRIAL-STIM Randomized Controlled Trial. Neuromodulation 2021 Apr;24(3):459-470 [FREE Full text] [doi: 10.1111/ner.13316] [Medline: 33258531]
11. Huang KT, Hazzard MA, Babu R, Ugiliweneza B, Grossi PM, Huh BK, et al. Insurance disparities in the outcomes of spinal cord stimulation surgery. Neuromodulation 2013;16(5):428-34; discussion 434. [doi: 10.1111/ner.12059] [Medline: 23647668]
12. Babu R, Hazzard MA, Huang KT, Ugiliweneza B, Patil CG, Boakye M, et al. Outcomes of percutaneous and paddle lead implantation for spinal cord stimulation: a comparative analysis of complications, reoperation rates, and health-care costs. Neuromodulation 2013;16(5):418-26; discussion 426. [doi: 10.1111/ner.12065] [Medline: 23647789]
13. Zhao P, Yoo I, Lancey R, Varghese E. Mobile applications for pain management: an app analysis for clinical usage. BMC Med Inform Decis Mak 2019 May 30;19(1):106 [FREE Full text] [doi: 10.1186/s12911-019-0827-7] [Medline: 31146739]
14. Schoth DE, Radhakrishnan K, Liossi C. A systematic review with subset meta-analysis of studies exploring memory recall biases for pain-related information in adults with chronic pain. PR9 2020 Mar 31;5(2):e816. [doi: 10.1097/pr9.0000000000000816]
15. Daoust R, Sirois M, Lee J, Perry J, Griffith L, Worster A, et al. Painful Memories: Reliability of Pain Intensity Recall at 3 Months in Senior Patients. Pain Res Manag 2017;2017:5983721 [FREE Full text] [doi: 10.1155/2017/5983721] [Medline: 28260963]
16. Gendreau M, Hufford MR, Stone AA. Measuring clinical pain in chronic widespread pain: selected methodological issues. Best Practice & Research Clinical Rheumatology 2003 Aug;17(4):575-592. [doi: 10.1016/s1521-6942(03)00031-7]
17. Odonkor CA. Fantastic Four: Age, Spinal Cord StimulatorWaveform, Pain Localization and History of SpineSurgery Influence the Odds of Successful SpinalCord Stimulator Trial. Pain Phys 2020 Jan 15;1;23(1;1):E19-E30. [doi: 10.36076/ppj.2020/23/e19]
18. Prabhala T, Kumar V, Gruenthal E, Collison C, Prusik J, Owusu S, et al. Use of a Psychological Evaluation Tool as a Predictor of Spinal Cord Stimulation Outcomes. Neuromodulation 2019 Feb;22(2):194-199. [doi: 10.1111/ner.12884] [Medline: 30378737]
19. 2020 Digital Health Consumer Survey. Accenture. 2020. URL: https://www.accenture.com/_acnmedia/PDF-130/Accenture-2020-Digital-Health-Consumer-Survey-US.pdf [accessed 2022-01-20]
20. Aging in Place Report. Rock Health. URL: https://mailchi.mp/rockhealth/aginginplace [accessed 2022-01-20]
21. Fulmer T, Reuben DB, Auerbach J, Fick DM, Galambos C, Johnson KS. Actualizing Better Health And Health Care For Older Adults. Health Aff (Millwood) 2021 Feb;40(2):219-225. [doi: 10.1377/hlthaff.2020.01470] [Medline: 33476185]
22. Han Y, Lu Y, Wang D, Ran M, Ren Q, Xie D, et al. The Use of Remote Programming for Spinal Cord Stimulation for Patients With Chronic Pain During the COVID-19 Outbreak in China. Neuromodulation 2021 Apr;24(3):441-447 [FREE Full text] [doi: 10.1111/ner.13382] [Medline: 33751731]
23. Lu Y, Xie D, Zhang X, Dong S, Zhang H, Yu B, et al. Management of Intractable Pain in Patients With Implanted Spinal Cord Stimulation Devices During the COVID-19 Pandemic Using a Remote and Wireless Programming System. Front Neurosci 2020;14:594696 [FREE Full text] [doi: 10.3389/fnins.2020.594696] [Medline: 33363453]
24. Consumer Health Behavior and the COVID-19 pandemic: What we've learned. PwC Health Research Institute. 2021. URL: https://www.pwc.com/us/en/industries/health-industries/library/assets/hri-2021-consumer-survey-Insight-chartpack.pdf [accessed 2022-01-20]
25. Telehealth Use Among Older Adults Before and During COVID-19. National Poll on Healthy Aging, University of Michigan. 2020. URL: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/156253/0212_NPHA-telehealth-report-FINAL-08142020-v6-handle.pdf?sequence=4&isAllowed=y [accessed 2022-01-20]
26. Zhang C, Zhu K, Lin Z, Huang P, Pan Y, Sun B, et al. Utility of Deep Brain Stimulation Telemedicine for Patients With Movement Disorders During the COVID-19 Outbreak in China. Neuromodulation 2021 Feb;24(2):337-342 [FREE Full text] [doi: 10.1111/ner.13274] [Medline: 33006811]

27. Jitkritsadakul O, Rajalingam R, Toenjes C, Munhoz RP, Fasano A. Tele-health for patients with deep brain stimulation: The experience of the Ontario Telemedicine Network. Mov Disord 2018 Mar;33(3):491-492. [doi: 10.1002/mds.27230] [Medline: 29119600]
28. Chen Y, Hao H, Chen H, Li L. The study on a telemedicine interaction mode for Deep Brain Stimulation postoperative follow-up. 2015 Presented at: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 25-29 August 2015; Milan, Italy. [doi: 10.1109/embc.2015.7318331]
29. Khazen O, Rosoklija G, Custozzo A, Gillogly M, Bridger C, Hobson E, et al. Correlation Between Aspects of Perceived Patient Loneliness and Spinal Cord Stimulation Outcomes. Neuromodulation 2021 Jan;24(1):150-155. [doi: 10.1111/ner.13299] [Medline: 33119206]
30. Duarte RV. Analysis of Psychological CharacteristicsImpacting Spinal Cord Stimulation TreatmentOutcomes: A Prospective Assessment. Pain Phys 2015 May 15;3;18(3;5):E369-E377. [doi: 10.36076/ppj.2015/18/e369]
31. Murphy KR, Han JL, Hussaini SMQ, Yang S, Parente B, Xie J, et al. The Volume-Outcome Effect: Impact on Trial-to-Permanent Conversion Rates in Spinal Cord Stimulation. Neuromodulation 2017 Apr;20(3):256-262 [FREE Full text] [doi: 10.1111/ner.12526] [Medline: 27696607]

Abbreviations

EMA: ecological momentary assessment
HIPAA: health insurance portability and accountability act
SCS: spinal cord stimulation

Edited by A Kushniruk; submitted 22.11.21; peer-reviewed by R Jamison; comments to author 17.01.22; revised version received 11.02.22; accepted 14.02.22; published 23.03.22.

Please cite as:
Lee JM, Woon R, Ramsum M, Halperin DS, Jain R
User Engagement and Assessment of Treatment Effectiveness in Patients Using a Novel Digital mHealth App During Spinal Cord Stimulation Screening Trials
JMIR eHum Factors 2022;9(1):e35134
URL: https://humanfactors.jmir.org/2022/1/e35134
doi:10.2196/35134
PMID:35167484

Original Paper

User Experience, Engagement, and Popularity in Mental Health Apps: Secondary Analysis of App Analytics and Expert App Reviews

Benjamin T Kaveladze[1], MA; Akash R Wasil[2], MA; John B Bunyi[1], MA; Veronica Ramirez[1], BA; Stephen M Schueller[1,3], PhD

[1]Department of Psychological Science, University of California, Irvine, CA, United States
[2]Department of Psychology, University of Pennsylvania, Philadelphia, CA, United States
[3]Department of Informatics, University of California, Irvine, CA, United States

Corresponding Author:
Benjamin T Kaveladze, MA
Department of Psychological Science
University of California
4201 Social & Behavioral Sciences Gateway
Irvine, CA, 92697-7085
United States
Phone: 1 310 984 9318
Email: bkavelad@uci.edu

Abstract

Background: User experience and engagement are critical elements of mental health apps' abilities to support users. However, work examining the relationships among user experience, engagement, and popularity has been limited. Understanding how user experience relates to engagement with and popularity of mental health apps can demonstrate the relationship between subjective and objective measures of app use. In turn, this may inform efforts to develop more effective and appealing mental health apps and ensure that they reach wide audiences.

Objective: We aimed to examine the relationship among subjective measures of user experience, objective measures of popularity, and engagement in mental health apps.

Methods: We conducted a preregistered secondary data analysis in a sample of 56 mental health apps. To measure user experience, we used expert ratings on the Mobile App Rating Scale (MARS) and consumer ratings from the Apple App Store and Google Play. To measure engagement, we acquired estimates of monthly active users (MAU) and user retention. To measure app popularity, we used download count, total app revenue, and MAU again.

Results: MARS total score was moderately positively correlated with app-level revenue (Kendall rank [T]=0.30, *P*=.002), MAU (T=0.39, *P*<.001), and downloads (T=0.41, *P*<.001). However, the MARS total score and each of its subscales (Engagement, Functionality, Aesthetics, and Information) showed extremely small correlations with user retention 1, 7, and 30 days after downloading. Furthermore, the total MARS score only correlated with app store rating at T=0.12, which, at *P*=.20, did not meet our threshold for significance.

Conclusions: More popular mental health apps receive better ratings of user experience than less popular ones. However, user experience does not predict sustained engagement with mental health apps. Thus, mental health app developers and evaluators need to better understand user experience and engagement, as well as to define sustained engagement, what leads to it, and how to create products that achieve it. This understanding might be supported by better collaboration between industry and academic teams to advance a science of engagement.

(JMIR Hum Factors 2022;9(1):e30766) doi:10.2196/30766

KEYWORDS

mental health apps; engagement; user experience; digital mental health; user retention

Introduction

Background

An increasing number of mental health apps are available to consumers, with estimates that 10,000 to 20,000 mental health apps currently exist [1,2]. Evidence suggests that these apps can help address various mental health concerns such as stress, depression, and anxiety. Even unguided apps intended for self-management can lead to reliable, albeit small, benefits [3], particularly for people with lower symptom severity [4]. The biggest challenge facing these apps, especially when provided in unguided, direct-to-consumer models, is engagement. Engagement with most mental health apps is abysmal—estimates suggest that most publicly available mental health apps for depression and anxiety have zero or near-zero active users [5]. This study aims to better understand engagement with mental health apps from the vantage points of user experience and popularity.

Previous Work on Mental Health App Engagement

Previous research has examined mental health app engagement in multiple ways [6]. A scoping review of concepts and components of engagement used in the digital health literature emphasized that engagement is a multifaceted concept with behavioral, cognitive, and affective components [7]. Although self-report engagement measures aim to capture some of these components [8], in this paper we focus on the behavioral component using analytic data, which tracks users' actual usage of apps (eg, number of downloads or average time per use). However, analytic data can also be used to determine different conceptualizations of engagement, and we consider some of these approaches below.

One approach to studying engagement is to quantify user retention: the proportion of users who continue to use an app over a certain period. Estimates suggest that approximately 4% of users who download a mental health app continue using it after 15 days, and 3% continue after 30 days [9]. Early efforts have also identified some factors that predict retention, such as therapeutic persuasiveness and therapeutic alliance [10]. Retention is a useful metric of engagement because it considers sustained use rather than only initial adoption (downloading an app).

Another perspective on engagement focuses on monthly active users (MAU): the number of people who use an app in a given month. Most mental health apps have nearly no active users, while a few apps have millions of active users [5]. This trend of vastly unequal distributions of users across similar apps seems to be true not only for mental health apps but also for apps focused on physical fitness and mood-tracking [11]. Given these extreme differences in MAU, efforts to understand how highly popular apps differ from unpopular apps have been a priority. Importantly, an app's MAU reflects two distinct components: (1) the number of people who downloaded the app (which reflects an app's popularity and marketing success) and (2) retention (which may reflect content and features within the app). Because retention data are often difficult to obtain, investigators recently proposed an alternative "stickiness" metric, defined as the number of monthly active users per normalized total downloads [12]. Interestingly, some of the most downloaded apps do not appear to be particularly sticky, and some of the stickiest apps are not the most downloaded [12].

There have also been efforts to improve engagement with digital mental health interventions. While this work is in its early stages, some promising strategies include incorporating human-centered design principles [13], branding digital mental health interventions in ways that appeal to specific subgroups of users [14], sending reminders and "digital triggers" [15], and incorporating human support [16].

In summary, current research has examined engagement from multiple perspectives, several efforts to improve engagement are underway, and research on engagement is still in its infancy. One important next step involves understanding why some apps are more engaging or more popular than others. Such work could inform efforts to increase engagement by highlighting specific content, features, characteristics, or development strategies that may contribute to engagement.

Characteristics of User Experience

User experience refers to the holistic experience of using a product such as a mobile app. It is shaped by an app's content, its functionality, and its look and feel. Similar to conceptualizations of engagement, it also encompasses affective, behavioral, and cognitive reactions and includes emotional, hedonic, and aesthetic variables [17]. User experience can be understood through various methods including expert or heuristic evaluations, user interviews, and user reviews [18]. As user experience is a multifaceted and complex concept, different methods of understanding user experience have relative strengths and weaknesses.

For mobile mental health apps, the most widely used measure of user experience is the mobile app rating scale (MARS) [19]. It has been used in various evaluations of health apps, including mindfulness apps [20] and pain-management apps [21]. The MARS evaluates mobile health app quality along dimensions of engagement, functionality, aesthetics, and information quality. The engagement scale assesses how interactive and interesting the app is, the functionality scale assesses the app's functioning and ease of use, the aesthetics scale assesses overall visual appeal and stylistic consistency, and the information subscale assesses the quality of the content. Averaged together, the 4 subscales form the MARS total score, which measures overall app quality. Typically, the MARS is used by a trained evaluator with expertise in some facet of mobile health apps such as technical or clinical expertise or lived experience with the health condition. In this way, the MARS can be thought of as a form of heuristic evaluation where experts score various components of the app using validated metrics. The MARS's construct validity was established by confirming its factor structure, and its concurrent validity was established by relating it to another app quality assessment tool [22]; however, more research is needed to determine the MARS's other psychometric properties, such as its predictive validity.

Another way to understand the user experience of mental health apps is to ask consumers, either directly through user experience

interviews or indirectly by analyzing consumer reviews posted to app stores [23]. Interviews can provide in-depth information but are labor-intensive and may not accurately reflect user behavior. App store reviews are easily accessible and plentiful for popular apps, but review-writers' perspectives may not be representative of most of an app's users. Studies that directly ask consumers about their experiences and those that leverage app store reviews can provide converging evidence of characteristics that are important for consumers such as a positive framing or simplicity [24,25]. A review of studies of mental health app user experience identified six themes among consumers' perceptions of apps: helpfulness, enhancements, technical issues, ease of use, satisfaction, and perceived issues [26]. Additionally, a study examining over 13,000 reviews of 106 mental health apps noted that user interface and user-friendliness were two of the most common aspects commented on by users and that poor usability was often noted as a reason for abandoning apps [23]. Although these themes align with some aspects of the MARS subscales, such as functionality, they also tend to correspond to more general perceptions of quality or specific improvements or deficits.

To summarize, methods for evaluating user experience for mobile health apps have been refined over the past years and produced useful insights into consumer preferences. Nonetheless, better understanding user experiences is critical because, ultimately, mental health apps are beneficial only insofar as users meaningfully engage with them.

This Study

This study aimed to identify associations between mental health app user experience and metrics of app popularity and engagement. We first hypothesized that more popular apps, in terms of app-level revenue, monthly active users, and downloads, would have higher user experience ratings. Second, we hypothesized that apps' levels of engagement, functionality, and aesthetic appeal would predict user retention more strongly than their informational quality. Third, we hypothesized that app store ratings would be correlated with user experience ratings.

Methods

Design and Material

We obtained MARS scores from One Mind PsyberGuide, a nonprofit organization that provides structured reviews of mental health apps [27]. One Mind PsyberGuide reviews mental health apps on multiple metrics including user experience as defined by the MARS. Three reviewers with training on MARS administration—2 PhD-level reviewers, each with extensive experience in user experience and mental health app reviews, and 1 individual with lived experience of mental health issues—completed each MARS review. These 3 ratings were averaged to produce the MARS scores provided on One Mind PsyberGuide and used in our analyses. Overall, we had access to MARS ratings from 91 mental health apps, including total score and all 4 MARS subscales. The ratings were completed between March 2020 and December 2020.

We obtained analytic data from Apptopia, a company that aggregates data on various metrics of mobile app usage and popularity [28]. This analytic data included app-level data on MAU (ie, the number of users who opened the app at least once in the past 30 days), daily revenue (in US$), daily downloads, app store rating (1-5, ratings were obtained from Google Play and Apple App Store and the mean across stores was used when data from both stores were available), and user retention variables corresponding to 1, 2, 3, 4, 5, 6, 7, 14, and 30 days after downloading the app (with values 0-100 corresponding to the percentage of people who opened the app *n* days after downloading it). The MAU, daily revenue, and daily downloads variables had daily values for each day between February 8, 2020, and February 8, 2021, which is a 1-year period that overlaps with that of the MARS ratings. We transformed these daily values to a single value per app, computed as the variable's mean across all of the days of the month in which that app's MARS review was completed. We performed all analyses using these month-averaged data, rather than the daily values. Because mean values as a measure of central tendency are susceptible to influence by outliers or skewness of the distribution, we also computed the variable's median values across all of the days of the month and report all analyses using these month median data in Multimedia Appendix 1. The app store rating variable and each retention variable had only one value in our data set. For app store rating, the value corresponded to the average of all app store ratings in the 365 days preceding February 8, 2021. Each retention variable's value reflected the average of retention values across every day in January 2021.

Exclusion Criteria

We chose to exclude several apps from analyses owing to missing data. First, we excluded apps that lacked Apptopia data for at least 1 day in the month as missing data may have created a bias toward inflated monthly average values. Among the 91 apps with MARS rating data provided by One Mind PsyberGuide, 56 had Apptopia data for every day of the month that the MARS review was completed, and 54 of those 56 had data for user retention and app store ratings. In total, 18 of the 91 apps with MARS rating data had no Apptopia data whatsoever. Thus, for some analyses we include 56 apps and for others we include 54 apps.

Data Availability

Our hypotheses and analysis plan were preregistered and are available on the internet, as well as the data sets we used for analyses (in addition to data on additional variables for each app) [29]. We have also provided the output from the main and sensitivity analyses in Multimedia Appendix 1. Owing to Apptopia's data-sharing policy, we have provided the Apptopia data separately with the app names deleted and the apps presented in random order.

Analyses

Statistical analyses were performed using the stats package in R (R Core Team, 2020). For all analyses, statistical significance was set at a preregistered threshold $P<.05$; however, we have reported exact P values unless they were <.001. Data

manipulation and figure creation were conducted in R using the tidyverse family of packages [30] and sjPlot [31].

For our first hypothesis—that more popular apps, in terms of app-level revenue, MAU, and downloads, would have higher MARS scores—we determined the Kendall rank (T) correlation coefficients (3 in total) between the MARS total score and revenue, MAU, and downloads. We used the Kendall rank correlation coefficient rather than the Pearson correlation coefficient because rank-order correlation is not overly impacted by the presence of extreme outliers (and we knew there were several such outliers in the revenue, MAU, and downloads variables in our data set) and is therefore more consistent with our research question. Nonetheless, this skewness remains important to consider when interpreting our results.

For our second hypothesis—that the MARS engagement, functionality, and aesthetic subscales would predict user retention more strongly than the information subscale—we chose to calculate Kendall rank correlation coefficients (12 in total) between each MARS subscale of interest (Engagement, Functionality, Aesthetics, and Information) and user retention, as measured by the percentage of users who downloaded the app who opened it 1, 7, and 30 days after download.

For our third hypothesis—that app store ratings would be correlated with the MARS total score—we chose to calculate the Kendall rank correlation coefficient between the MARS total score and app store rating.

Results

Results Overview

Figure 1 shows variable distributions and mean values, and Figure 2 illustrates Kendall rank correlation coefficients across variables of interest. MARS total scores tended to be high (mean 3.85, SD 0.65), as did the MARS subscales for engagement (mean 3.68, SD 0.76), functionality (mean 4.16, SD 0.61), aesthetics (mean 3.84, SD 0.84), and information (mean 3.72, SD 0.69). App store ratings were also high (mean 4.39, SD 0.59). Revenue, MAU, and downloads were highly skewed owing to a few extremely popular outliers.

Figure 1. Frequency distributions for all variables (n=54 or n=56) included in our analyses. Variable mean values are shown with gold vertical lines. MARS: Mobile App Rating Scale.

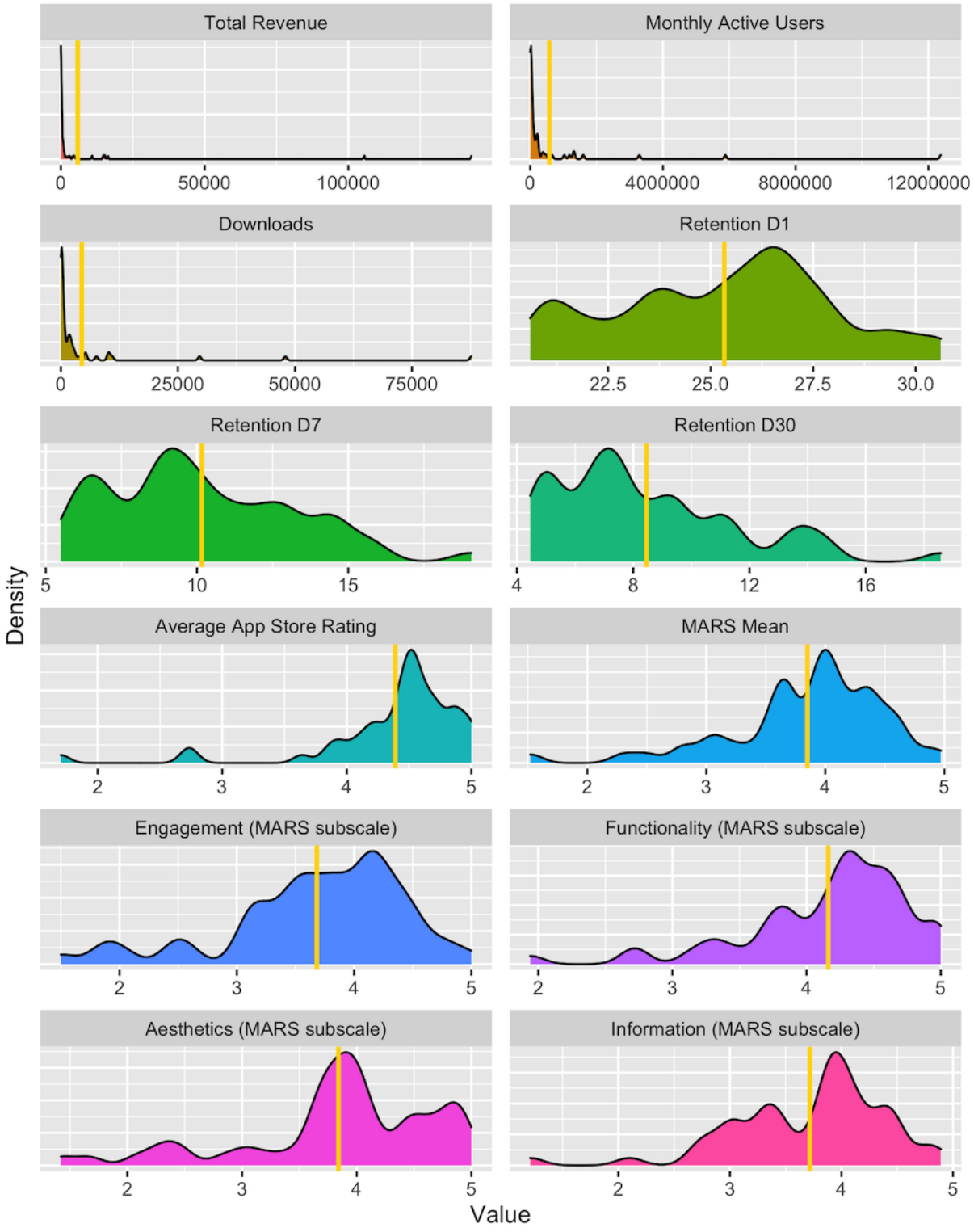
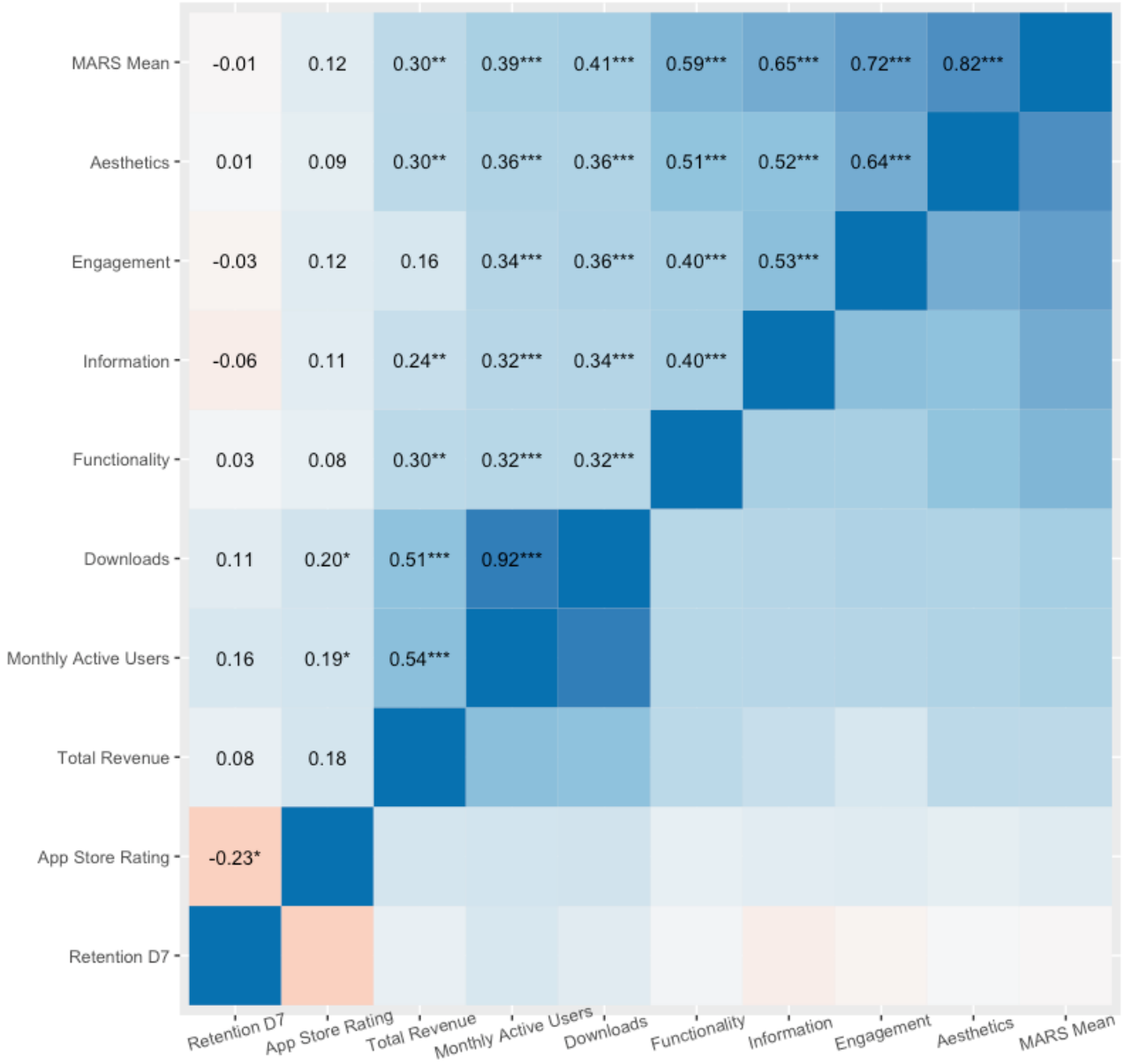


Figure 2. Kendall rank correlation coefficients between all variables (n=54 or n=56). Pairwise deletion was used to deal with missing data. *P<.05. **P<.01. ***P<.001. MARS: Mobile App Rating Scale.



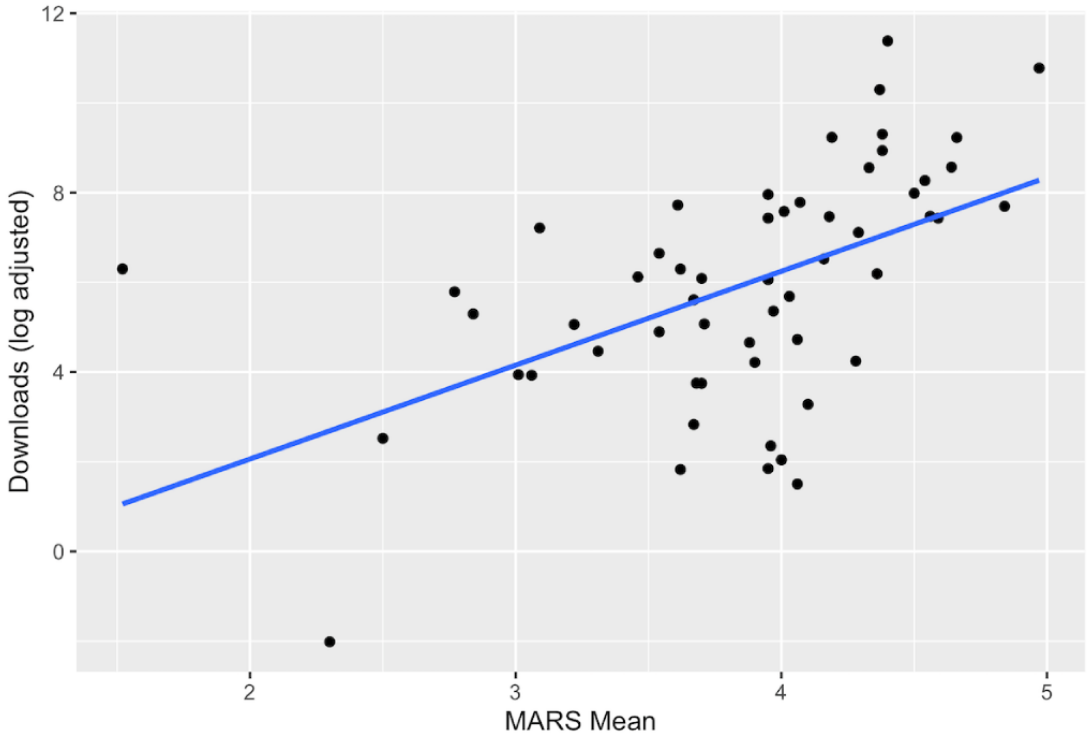
Distribution of App Usage

The distribution of monthly active users among the 56 mental health apps we examined was highly skewed (mean 578,645, SD 1,856,468, median 48,676, range 0-12,373,122). Among all the apps in our data set, the 3 most popular apps accounted for 66.6% of MAUs, and the 10 most popular apps accounted for 90% of MAUs [11]. As noted in the analyses section, we account for these distributions' skewness by using Kendall rank correlations in our analyses.

Associations Between MARS Scores and App Popularity, User Retention, and App Store Ratings

In our sample of 56 mental health apps, Kendall rank correlation analyses revealed that the MARS total score was moderately positively correlated with app-level revenue (T=0.30, P=.002), MAU (T=0.39, P<.001), and downloads (T=0.41, P<.001; Figure 3). Conversely, in our sample of 54 mental health apps, the MARS total score and its subscales (ie, Engagement, Functionality, Aesthetics, and Information) showed minor associations with user retention 1, 7, and 30 days after downloading (T=–0.10 to 0.17) and none of these associations met our threshold for significance (P=.93 to .07). Lastly, the MARS total score was also extremely weakly correlated with app store ratings (T=0.12), which did not meet our threshold for significance (P=.20).

Figure 3. The association between Mobile App Rating Scale (MARS) Mean scores and log-adjusted downloads for mobile mental health apps, Kendall T=0.41, *P*<.001, n=56.



Sensitivity Analysis

The daily values that were averaged to form the monthly average values for the MAU, daily revenue, and daily downloads variables were not normally distributed for many apps, although they also did not have extreme outliers. As sensitivity analysis, we reran the analyses using the median of the MAU, daily revenue, and daily downloads variables instead of the mean. The results from the two approaches were nearly identical, and we have included the output from the analyses using median values in Multimedia Appendix 1.

Discussion

Principal Findings

Our findings support one of our three hypotheses. Specifically, we found that user experience scores were related to several app popularity metrics: downloads, revenue, and monthly active users. However, none of the MARS subscales were predictive of user retention. Therefore, user experience, at least as defined by the MARS, is a fairly good indicator of how many people might start to engage with, or adopt, an app but may be less informative about users' sustained engagement with the app. Further, the lack of a correlation between user experience (as measured by the MARS total score) and app store ratings suggests that app store ratings are not a broadly useful measure of user experience.

Relationship Between Popularity Metrics and User Experience

The moderate rank-order correlations we observed between popularity or revenue and user experience suggest that higher scores on user experience, as rated by individual observers, are characteristic of more popular apps. Importantly, however, this relationship does not appear linear: a few apps are responsible for nearly all users, suggesting that marginal improvements to popularity may not be sufficient to retain users. Instead, reaching a "popularity threshold" may be necessary, with apps below that threshold being unlikely to gain many users.

These findings suggest that industry teams, rather than academic ones, may be best-suited to create highly engaging and popular products. To date, nearly all of the most engaging digital mental health interventions have been developed by industry teams, rather than academic teams [12,32]. It is plausible that differences in funding sources (eg, federal grants vs flexible capital), incentive structures (eg, priorities on publishing vs marketing), timelines (eg, multi-year studies vs rapid testing of prototypes), and other factors may give industry teams a competitive advantage in developing highly engaging products. As an example, Headspace and Calm, the two most popular mental health and wellness apps, each raised over US $140 million in funding from venture capital firms [33]. Furthermore, industry teams are often diverse and interdisciplinary and charged not only with developing an engaging product but also marketing and financing that product. Aspects of business models and marketing may play an important role in people's likelihood to adopt or sustain use of a given app. For example, word of mouth is a common way that people learn about mental health apps [1]. Payment models also impact both adoption and sustainment; people prefer free apps [1] but dislike "freemium pricing" [34]. Some mental health apps have also used celebrity advertising, such as LeBron James for Calm or Michael Phelps for Talkspace. This paper focused on aspects of the apps themselves rather than these other aspects of the business models and advertising, but those aspects are worth exploring in further work.

The Need to Better Advance a Science of Engagement

A common presumption is that mental health apps require sustained engagement for users to experience their intended benefits. Taking a similar approach to the National Institute of Mental Health Research Domain Criteria framework, Graham et al [35] propose that engagement is a critical mechanism of action for mental health apps. In their conceptualization, engagement can be separated into elements that speak to design targets such as usefulness, usability, and satisfaction, as well as use metrics, such as those used in our analyses. Our findings suggest, however, that use metrics might speak to different aspects of engagement, and that more work needs to unpack elements that lead people to adopt, use, and explore technologies. This is logically similar to efforts in implementation science frameworks to map key implementation outcomes including adoption, appropriateness, and sustainment, and consistent with work that has mapped those outcomes onto different variables relevant to technology-enabled services [36].

As can be seen in Figure 2, while our measure of initial engagement—ie, downloads—was predicted by several variables, our measure of sustained engagement—ie, retention—was not. While retention was quite low in general, it still varied considerably across apps (retention 7 days after download ranged from 5.5% to 19.1%). Our study echoes previous research suggesting that retention is particularly challenging to understand [9]. One reason that retention may be so challenging to predict using the statistical approaches typical in behavioral sciences is that on the user level, retention's distribution is highly skewed: most users do not engage with a given app more than once or twice but some engage much more often; hence, mean retention values do not represent most users' experiences. More in-depth forms of analysis such as user interviews and longitudinal analyses may be required to understand patterns in mental health app user retention. These qualitative approaches can go beyond just how many people use or stop using an app to explore richer questions regarding users' journeys with an app. In turn, these data can help to identify critical aspects of the user experience.

It is also worth noting that user engagement, even retention, is likely a heterogeneous concept. Even among users who are considered to be retained, patterns of use might differ, including among different dimensions such as frequency (as in consistent vs bursty use), intensity (as in moderate or super users [37]), time (as in circadian patterns in use [38]), or type (as in using clinically meaningful app features [39]). These dimensions similarly characterize other types of complex behavior such as exercise (ie, the Frequency, Intensity, Time, and Type model). For some users, decreased use over time could be a sign that the app improved their well-being or that the user completed the app's intervention as intended. Therefore, although much has been made of the poor rates of long-term sustainment in mental health apps [9-11], some users likely experience "happy abandonment," wherein a lack of sustained use suggests they received what they needed. Unfortunately, given that our user retention information was obtained from app-level analytic data, we were not able to determine individual-level characteristics of retention and engagement. However, future work could help determine the degree to which app engagement patterns are shaped by characteristics of the app, such as user experience or app features, and characteristics of the different ways that people use digital health products.

Another reason that retention might be a heterogeneous concept is that mental health apps vary in their intended user journeys. Some apps might be designed for people to use them every day, whereas others might be designed for more emergent yet infrequent situations. Therefore, in addition to individual-level characteristics of retention and engagement, it is also worth noting that retention might have app-level characteristics; as such, retention may not allow apples-to-apples comparisons of apps. Again, because our data were obtained from an analytics platform rather than the apps themselves, we were not able to conduct more nuanced analyses of retention; however, efforts that could combine and synthesize engagement data across platforms [40] could help investigate these questions among others.

Given that achieving sustained engagement is so difficult for mental health apps, an alternative strategy involves circumventing the challenge of long-term engagement altogether by creating digital interventions that are designed to confer benefit rapidly. An example comes from the growing literature on digital single-session interventions, which are designed to produce benefits after just one sitting [41-43]. These interventions attempt to reimagine how to support users' mental health in ways that differ from typical therapist-client interactions but might hold greater appeal and utility.

Limitations

There are several limitations to consider in this work. First, the set of 56 apps observed in this study (for which One Mind PsyberGuide chose to complete MARS reviews and for which engagement data were available from Apptopia) are not representative of the full array of available mental health apps, with a likely bias toward more popular apps and those designed for English-speaking audiences. Nonetheless, these apps represented a fairly wide range of values across all variables. Second, the MARS may not be an ideal measure of user experience. Although many of the elements in the MARS address user engagement, it is often conceptualized as an overall measure of app quality, rather than solely user experience [19]. Furthermore, some aspects of user experience within apps might not be captured by the MARS, such as gamification principles [44]. Third, because the study is cross-sectional and observational, we are unable to infer causality. Many of the observed relationships between variables are likely bidirectional; for example, better user experience likely causes apps to become more popular, but apps that are more popular also gain the resources to improve their user experience design. Fourth, apps differed in the time distances between their respective MARS review dates and the dates for which their rating and retention data were available, although all MARS ratings occurred during the 1-year period for which analytic data were obtained. Lastly, as the aim of this paper was to understand how aspects of user experience relate to engagement and popularity, we do not know if using these mental health apps actually helps people to achieve their goals in using these apps or to derive clinical benefits. In this study, we did not have access to analytic data on user

outcomes, but such data would be a strength of a solution that facilitates better collaboration with developers for analysis and evaluation purposes.

Future Directions

Although this analysis observes rank-order trends, it does not explain why a few apps, such as Calm and Headspace, are exponentially more popular than others. Future research can explore the complex combination of factors, such as marketing dollars and market trends, which could explain these few apps' outsized popularity. Such research might also explore the optimal conditions for making influential and effective apps. For example, industry teams tend to create more popular and engaging apps than academic teams do; however, solutions to user engagement problems plaguing these apps might be best pursued by rigorous research combining quantitative, qualitative, and experimental approaches [45]. Thus, research could examine if collaborations between academic and industry teams may be particularly fruitful in creating evidence-based and highly scalable interventions. Finally, given low retention among mental health apps, future work should explore innovative intervention strategies by which apps can support mental health in ways that appeal to users.

As one resource for exploring these future directions and other ideas, we encourage researchers to explore the publicly available data sets that we used to conduct our analyses [29], which contain data on more apps and more variables (including credibility, intervention target, intervention approach, app price, and average time spent per session) than those examined in this study.

Conclusions

We found that popular mental health apps—as defined by their number of downloads, revenue, and number of monthly active users—tend to be rated as having a better user experience than less popular apps. We also noted that user retention metrics are not well-predicted by other app-level metrics. We encourage further collaboration between industry and academic teams to better advance a science of engagement and to create more effective and appealing mental health apps.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preregistered analysis and sensitivity analysis output.
[PDF File (Adobe PDF File), 1155 KB - humanfactors_v9i1e30766_app1.pdf]

References

1. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and Interest in Health Apps Among Those With Mental Health Needs: Survey and Focus Group Study. J Med Internet Res 2018 Jun 11;20(6):e10141 [FREE Full text] [doi: 10.2196/10141] [Medline: 29891468]
2. Torous J, Roberts LW. Needed Innovation in Digital Health and Smartphone Applications for Mental Health: Transparency and Trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
3. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
4. Karyotaki E, Efthimiou O, Miguel C, Bermpohl FMG, Furukawa TA, Cuijpers P, Individual Patient Data Meta-Analyses for Depression (IPDMA-DE) Collaboration, et al. Internet-Based Cognitive Behavioral Therapy for Depression: A Systematic Review and Individual Patient Data Network Meta-analysis. JAMA Psychiatry 2021 Apr 01;78(4):361-371. [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
5. Wasil AR, Gillespie S, Shingleton R, Wilks CR, Weisz JR. Examining the Reach of Smartphone Apps for Depression and Anxiety. Am J Psychiatry 2020 May 01;177(5):464-465. [doi: 10.1176/appi.ajp.2019.19090905] [Medline: 32354266]
6. Ng MM, Firth J, Minen M, Torous J. User Engagement in Mental Health Apps: A Review of Measurement, Reporting, and Validity. Psychiatr Serv 2019 Jul 01;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
7. Kelders SM, van Zyl LE, Ludden GDS. The Concept and Components of Engagement in Different Domains Applied to eHealth: A Systematic Scoping Review. Front Psychol 2020;11:926 [FREE Full text] [doi: 10.3389/fpsyg.2020.00926] [Medline: 32536888]
8. Kelders SM, Kip H, Greeff J. Psychometric Evaluation of the TWente Engagement with Ehealth Technologies Scale (TWEETS): Evaluation Study. J Med Internet Res 2020 Oct 09;22(10):e17757 [FREE Full text] [doi: 10.2196/17757] [Medline: 33021487]
9. Baumel A, Muench F, Edan S, Kane JM. Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]

10. Baumel A, Kane JM. Examining Predictors of Real-World User Engagement with Self-Guided eHealth Interventions: Analysis of Mobile Apps and Websites Using a Novel Dataset. J Med Internet Res 2018 Dec 14;20(12):e11491 [FREE Full text] [doi: 10.2196/11491] [Medline: 30552077]
11. Wasil AR, Gillespie S, Schell T, Lorenzo-Luaces L, DeRubeis RJ. Estimating the real-world usage of mobile apps for mental health: development and application of two novel metrics. World Psychiatry 2021 Feb;20(1):137-138 [FREE Full text] [doi: 10.1002/wps.20827] [Medline: 33432761]
12. Carlo AD, Hosseini Ghomi R, Renn BN, Strong MA, Areán PA. Assessment of Real-World Use of Behavioral Health Mobile Applications by a Novel Stickiness Metric. JAMA Netw Open 2020 Aug 03;3(8):e2011978 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.11978] [Medline: 32744628]
13. Lyon AR, Koerner K. User-Centered Design for Psychosocial Intervention Development and Implementation. Clin Psychol (New York) 2016 Jun;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
14. Rith-Najarian L, Sun W, Chen A, Chorpita B, Chavira D, Mougalian S, et al. What's in a name? Branding of online mental health programming for university students. J Consult Clin Psychol 2019 Apr;87(4):380-391. [doi: 10.1037/ccp0000383] [Medline: 30883165]
15. Muench F, Baumel A. More Than a Text Message: Dismantling Digital Triggers to Curate Behavior Change in Patient-Centered Health Interventions. J Med Internet Res 2017 May 26;19(5):e147 [FREE Full text] [doi: 10.2196/jmir.7463] [Medline: 28550001]
16. 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]
17. Law E, Roto V, Hassenzahl M, Vermeeren A, Kort J. Understanding, scoping and defining user experience: a survey approach. 2009 Presented at: CHI '09: CHI Conference on Human Factors in Computing Systems; April 4-9, 2009; Boston, MA. [doi: 10.1145/1518701.1518813]
18. Lemon C, Huckvale K, Carswell K, Torous J. A Narrative Review of Methods for Applying User Experience in the Design and Assessment of Mental Health Smartphone Interventions. Int J Technol Assess Health Care 2020;36(1):64-70. [doi: 10.1017/S0266462319003507] [Medline: 31973787]
19. 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]
20. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and Evaluation of Mindfulness-Based iPhone Apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
21. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the Quality of Mobile Apps for the Management of Pain: Systematic Search and Evaluation Using the Mobile App Rating Scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
22. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). PLoS One 2020;15(11):e0241480 [FREE Full text] [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
23. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. Health Informatics J 2020 Sep;26(3):2042-2066 [FREE Full text] [doi: 10.1177/1460458219896492] [Medline: 31920160]
24. Schueller SM, Armstrong CM, Neary M, Ciulla RP. An Introduction to Core Competencies for the Use of Mobile Apps in Cognitive and Behavioral Practice. Cognitive and Behavioral Practice 2021 Jan. [doi: 10.1016/j.cbpra.2020.11.002]
25. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User Experience of Cognitive Behavioral Therapy Apps for Depression: An Analysis of App Functionality and User Reviews. J Med Internet Res 2018 Jun 06;20(6):e10120 [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]
26. Chan AHY, Honey MLL. User perceptions of mobile digital apps for mental health: Acceptability and usability - An integrative review. J Psychiatr Ment Health Nurs 2022 Feb;29(1):147-168. [doi: 10.1111/jpm.12744] [Medline: 33604946]
27. Apps and digital health resources reviewed by experts. One Mind PsyberGuide. URL: https://onemindpsyberguide.org/ [accessed 2021-11-14]
28. Outsmart Your Competition. Apptopia. URL: https://apptopia.com/ [accessed 2021-11-14]
29. Kaveladze B, Bunyi J, Ramirez V, Schueller S, Wasil A. Relating mental health apps' MARS ratings to app-level metrics. OSF. 2021 Feb 01. URL: https://osf.io/rhgy7/ [accessed 2021-07-29]
30. Wickham H. tidyverse: Easily Install and Load the 'Tidyverse'. RStudio. URL: https://CRAN.R-project.org/package=tidyverse [accessed 2021-07-15]
31. Lüdecke D, Bartel A, Schwemmer C, Powell C, Djalovski A, Titz J. sjPlot: Data Visualization for Statistics in Social Science Internet. RStudio. URL: https://CRAN.R-project.org/package=sjPlot [accessed 2021-07-15]
32. Wasil AR, Gillespie S, Patel R, Petre A, Venturo-Conerly KE, Shingleton RM, et al. Reassessing evidence-based content in popular smartphone apps for depression and anxiety: Developing and applying user-adjusted analyses. J Consult Clin Psychol 2020 Nov;88(11):983-993. [doi: 10.1037/ccp0000604] [Medline: 32881542]
33. Shah RN, Berry OO. The Rise of Venture Capital Investing in Mental Health. JAMA Psychiatry 2020 Sep 16. [doi: 10.1001/jamapsychiatry.2020.2847] [Medline: 32936238]

34. Chiauzzi E, Newell A. Mental Health Apps in Psychiatric Treatment: A Patient Perspective on Real World Technology Usage. JMIR Ment Health 2019 Apr 22;6(4):e12292 [FREE Full text] [doi: 10.2196/12292] [Medline: 31008711]
35. Graham AK, Kwasny MJ, Lattie EG, Greene CJ, Gupta NV, Reddy M, et al. Targeting subjective engagement in experimental therapeutics for digital mental health interventions. Internet Interv 2021 Sep;25:100403 [FREE Full text] [doi: 10.1016/j.invent.2021.100403] [Medline: 34401363]
36. Hermes ED, Lyon AR, Schueller SM, Glass JE. Measuring the Implementation of Behavioral Intervention Technologies: Recharacterization of Established Outcomes. J Med Internet Res 2019 Jan 25;21(1):e11752 [FREE Full text] [doi: 10.2196/11752] [Medline: 30681966]
37. Sanatkar S, Baldwin PA, Huckvale K, Clarke J, Christensen H, Harvey S, et al. Using Cluster Analysis to Explore Engagement and e-Attainment as Emergent Behavior in Electronic Mental Health. J Med Internet Res 2019 Nov 28;21(11):e14728 [FREE Full text] [doi: 10.2196/14728] [Medline: 31778115]
38. Theilig M, Knapp A, Nicholas J, Zarnekow R, Mohr D. Characteristic Latent Features for Analyzing Digital Mental Health Interaction and Improved Explainability. 2020 Sep 2020:1. [doi: 10.21203/rs.3.rs-68076/v1]
39. Zhang R, Nicholas J, Knapp AA, Graham AK, Gray E, Kwasny MJ, et al. Clinically Meaningful Use of Mental Health Apps and its Effects on Depression: Mixed Methods Study. J Med Internet Res 2019 Dec 20;21(12):e15644 [FREE Full text] [doi: 10.2196/15644] [Medline: 31859682]
40. Powell AC, Neary M, Schueller SM. Mental Health Apps: Ensuring Quality and Reimbursement Through a Dynamic Payment Formulary. Psychiatr Serv 2021 May 01;72(5):614. [doi: 10.1176/appi.ps.202000754] [Medline: 33950744]
41. Osborn TL, Rodriguez M, Wasil AR, Venturo-Conerly KE, Gan J, Alemu RG, et al. Single-session digital intervention for adolescent depression, anxiety, and well-being: Outcomes of a randomized controlled trial with Kenyan adolescents. J Consult Clin Psychol 2020 Jul;88(7):657-668. [doi: 10.1037/ccp0000505] [Medline: 32391709]
42. Schleider J, Weisz J. A single-session growth mindset intervention for adolescent anxiety and depression: 9-month outcomes of a randomized trial. J Child Psychol Psychiatry 2018 Feb;59(2):160-170. [doi: 10.1111/jcpp.12811] [Medline: 28921523]
43. Wasil AR, Taylor ME, Franzen RE, Steinberg JS, DeRubeis RJ. Promoting Graduate Student Mental Health During COVID-19: Acceptability, Feasibility, and Perceived Utility of an Online Single-Session Intervention. Front Psychol 2021;12:569785 [FREE Full text] [doi: 10.3389/fpsyg.2021.569785] [Medline: 33897511]
44. Floryan M, Chow PI, Schueller SM, Ritterband LM. The Model of Gamification Principles for Digital Health Interventions: Evaluation of Validity and Potential Utility. J Med Internet Res 2020 Jun 10;22(6):e16506 [FREE Full text] [doi: 10.2196/16506] [Medline: 32519965]
45. Mohr DC, Schueller SM, Riley WT, Brown CH, Cuijpers P, Duan N, et al. Trials of Intervention Principles: Evaluation Methods for Evolving Behavioral Intervention Technologies. J Med Internet Res 2015 Jul 08;17(7):e166 [FREE Full text] [doi: 10.2196/jmir.4391] [Medline: 26155878]

Abbreviations

MARS: Mobile App Rating Scale
MAU: monthly active users

Edited by A Kushniruk; submitted 09.08.21; peer-reviewed by A Rovetta, C Moberg; comments to author 06.11.21; revised version received 24.11.21; accepted 27.11.21; published 31.01.22.

Please cite as:
Kaveladze BT, Wasil AR, Bunyi JB, Ramirez V, Schueller SM
User Experience, Engagement, and Popularity in Mental Health Apps: Secondary Analysis of App Analytics and Expert App Reviews
JMIR Hum Factors 2022;9(1):e30766
URL: https://humanfactors.jmir.org/2022/1/e30766
doi:10.2196/30766
PMID:35099398

Review

Boundary Objects as Dialogical Learning Accelerators for Social Change in Design for Health: Systematic Review

Gijs Terlouw[1,2], MSc; Derek Kuipers[3], PhD; Lars Veldmeijer[1], MA; Job van 't Veer[4], PhD; Jelle Prins[5], PhD; Jean-Pierre Pierie[6,7], MD, PhD

[1]NHL Stenden University of Applied Sciences, Leeuwarden, Netherlands

[2]Medical Faculty Lifelong Learning, Education & Assessment Research Network, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

[3]Research Group Serious Gaming, NHL Stenden University of Applied Sciences, Leeuwarden, Netherlands

[4]Research Group Digital Innovation in Healthcare and Social Work, NHL Stenden University of Applied Sciences, Leeuwarden, Netherlands

[5]University Medical Center Groningen, University of Groningen, Groningen, Netherlands

[6]Post Graduate School of Medicine, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

[7]Department of Surgery, Medical Center Leeuwarden, Leeuwarden, Netherlands

Corresponding Author:
Gijs Terlouw, MSc
NHL Stenden University of Applied Sciences
Rengerslaan 10
Leeuwarden, 8919 DD
Netherlands
Phone: 31 628317457
Email: gijs.terlouw@nhlstenden.com

Abstract

Background: Boundary objects can add value for innovative design and implementation research in health care through their organizational focus and the dynamic structure between ill-structured and tailored use. However, when innovation is approached as a boundary object, more attention will need to be paid to the preimplementation phase. Research and design thinking pay attention to the preimplementation stage but do not have a social or organizational focus per se. The integration of boundary objects in design methodologies can provide a more social and organizational focus in innovative design projects by mapping out the mechanisms that occur at boundaries during design. Four dialogical learning mechanisms that can be triggered at boundaries have been described in the literature: identification, coordination, reflection, and transformation. These mechanisms seem suitable for integration in innovative design research on health.

Objective: Focusing on innovation in health, this study aims to find out whether the different learning mechanisms can be linked to studies on health innovation that mention boundary objects as a concept and assess whether the related mechanisms provide insight into the stage of the design and implementation or change process.

Methods: The following 6 databases were searched for relevant abstracts: PubMed, Scopus, Education Resources Information Center, PsycINFO, Information Science and Technology Abstracts, and Embase. These databases cover a wide range of published studies in the field of health.

Results: Our initial search yielded 3102 records; after removing the duplicates, 2186 (70.47%) records were screened on the title and abstract, and 25 (0.81%) papers were included; of the 13 papers where we identified 1 mechanism, 5 (38%) described an innovation or innovative project, and of the 12 papers where we identified more mechanisms, 9 (75%) described the development or implementation of an innovation. The reflective mechanism was not identified solely but was present in papers describing a more successful development or implementation project of innovation. In these papers, the predetermined goals were achieved, and the process of integration was relatively smoother.

Conclusions: The concept of boundary objects has found its way into health care. Although the idea of a boundary object was introduced to describe how specific artifacts can fulfill a bridging function between different sociocultural sites and thus have a social focus, the focus in the included papers was often on the boundary object itself rather than the social effect. The reflection and transformation mechanisms were underrepresented in the included studies but based on the findings in this review, pursuing

to trigger the reflective mechanism in design, development, and implementation projects can lead to a more fluid and smooth integration of innovation into practice.

(JMIR Hum Factors 2022;9(1):e31167) doi:10.2196/31167

KEYWORDS

boundary objects; health; innovation; design; systematic review

Introduction

Background

The concept of boundary objects was introduced in 1989 by Star [1] to describe how specific artifacts can fulfill a bridging function between different sociocultural sites. Over the past decades, there has been more interest in boundaries, boundary crossing, and boundary objects [2-6]. The idea of boundary objects was initially framed to facilitate constructive cooperation between sites or social systems without consensus [7]. This organizational feature of boundary objects can be of great value in health care innovation, where the promise of innovations outweighs their actual impact [8-13]. Owing to many different stakeholders and parties in health care with different needs and goals, the implementation of innovations in health care practice is complex [14]. Many frameworks on innovation and implementation pursue consensus [15-21], mainly from a monodisciplinary approach, where at some point, all parties and stakeholders must be convinced that an innovation is of added value from a specific viewpoint. Within disciplines, this can be feasible, but across disciplines, this is often challenging. Boundary objects offer a different perspective on this issue. Boundary objects ideally address the needs of each stakeholder group and aim to contribute to the goals of all stakeholders involved, even if they do not pursue the same goal. This also means that different stakeholders can interpret a boundary object differently, something that Star [7] calls *interpretive flexibility*. However, not striving for consensus but identifying and addressing needs on the front end requires a fundamentally different approach. Much more attention will need to be paid to the preimplementation phase, which is seldom included in frameworks [22]. The design discipline is a discipline par excellence that pays attention to precisely this phase.

Design research and design thinking are increasingly finding their way into the health care sector as appropriate methodologies of responding to a world with more open, complex, and increasingly networked problems. Design holds the promise of offering suitable strategies for complex problems and actively involves stakeholders during the development and implementation of innovation [23,24]. Design as a discipline already has a long history in the development of medical devices but is now broadening its scope in shaping the future of health care [25-28]. Owing to different causes, the worlds of health care and design are converging. In health care, there is a shift in focus toward patient experience and values, increasing the quality of life and patients' participation in care and treatment [29,30]. In the design discipline, developments toward phenomena such as experience design [31], value-sensitive design [32], and people's involvement in design through participatory design [33-35] seem to have a good fit with the shifts in the focus of health care. The focus of emerging design disciplines on innovation, transformation, and services within organizations [36] can also solve implementation and adoption problems in innovation. However, many frameworks or models that provide insight into shaping the design process focus more specifically on the steps, methods, or guide points essential for developing an artifact [37-42] and less on shaping the process of change.

Boundary Objects Within Design Research

Overview

The concept of boundary objects has also been applied in different studies on design and product development [43-45]. However, the focus is often too specific on 1 element of boundary objects: interpretive flexibility. Other elements mentioned by Star [7] are the structure of informatics, work process needs and arrangements, and the dynamic between the ill-structured and more tailored uses of the object [1,7,46]. In citations, the aspect of interpretive flexibility is overrepresented: "boundary objects almost became synonymous with interpretive flexibility" [7]. Nevertheless, in design, the interest in interpretative flexibility as a feature of boundary objects is sensible. By developing and testing concepts and prototypes with stakeholders, a lot can be learned regarding the product or idea during development, primarily through the interpretation of end users. However, the more organizational side of boundary objects to let people work together constructively and the focus of boundary objects in changing organizations can be of added value, especially in the embedding and adoption of innovation in practice. In the life cycle of boundary objects, Star [7] describes boundary objects' role in the organizational nature as a back-and-forth movement between ill-structured and well-structured [7]. In this life cycle, parallels can be drawn with an innovation or design life cycle, a back-and-forth movement from the emergence of a complex problem in an ill-structured context to the development of a solution that ideally leads to integration in the context where a new and clear structure occurs.

Both boundary objects—through their organizational focus and the dynamic structure between ill-structured and tailored use [7]—and design—which, in complex settings, requires managing and moving multiple stakeholders from the problem to the solution space [47]—might be able to guide innovative transformations in health care. However, frameworks aimed at adoption and implementation rarely pay attention to the development process [22]; frameworks aimed at design are often more product oriented.

A possible starting point to provide more social focus during the application of boundary objects in the development and transformation of innovation can be found in a systematic review

by Akkerman and Bakker [3], who described four dialogical learning mechanisms that can take place at boundaries: identification, coordination, reflection, and transformation. The mechanisms are similar to interrelational forms of boundary work that Langley et al [48] reported to describe organizational work: competitive boundary work can be linked to identification, collaborative boundary work can be related to coordination, and configurational boundary work seems synonymous with transformation. Although the initial focus of the 4 mechanisms of Akkerman and Bakker [3] was mainly on education, they seem to fit with well-known focus areas in design.

Identification

The identification mechanism is about learning what the diverse practices are to each other [3]. Typically, in identification processes, the boundaries are encountered, reconstructed, or reinforced. The identification mechanism is not necessarily related to overcoming discontinuities. The strategy to enact this process is something that designers often perform in the first phase of research. Many design research thinking or design research projects start with a phase or focus, such as *empathy* [49-51], *discover* [52-54] or *assess user needs, analyze content, and context* [55-57], in which, inter alia, user needs are identified. In design, this phase is essential. It allows designers and researchers to comprehend the situations and perspectives of others [58]. During this phase, methods such as *empathy maps* [59], *personas* [60,61], and *a day in the life* [62] are used to identify visible and invisible components that define the stakeholders' identities and needs, which fits the dialogical learning mechanism of identification. The stakeholders and designers are reinforced in their roles and (professional) identities; boundaries are encountered, reconstructed, or reinforced but not overcome in this phase.

Coordination

Dialogical learning mechanism coordination is mainly about creating cooperative and constructive exchanges between practices, even without consensus [3]. This description is the closest to the concept that Star [1,7,46] originally presented. It is crucial for design teams to work together effectively and constructively in the design discipline, even if the backgrounds and practices differ. Within the design, multilayered interactions can occur, and through the development of co-design practices, users can become active participants in design projects and processes [63,64]. There is a wide variety of methods in design to facilitate constructive collaboration between practices in health innovation, such as *hackathons* [65-67], *future workshops* [68,69], and other creative participatory design methods [70-72]. The potential of the coordination is in (temporarily) overcoming boundaries and getting to know each other, not in reconstructing them. This usually fits the design stage where there are no objectives formulated yet; the problem still needs to be defined, and the co-ownership of different stakeholders is desirable.

Reflection

The dialogical learning mechanism reflection emphasizes the role of boundary crossing and boundary objects in realizing, clarifying, and exchanging differences between practices [3]. Reflection is about expanding perspectives through perspective taking and perspective making. Together with the identification mechanism, the reflective mechanism focuses mainly on meaning-oriented learning processes. Once enacted, the reflective mechanism results in an expanded set of perspectives that inform future practices. Within the design discipline, and as a designer, reflection is essential. Schön [73] describes the creative process as a continuous process of reflection in action. Following the theory by Schön [73], designers enfold a continuum of activity by reflecting and acting within a new situation. The designer and stakeholder reflections help to frame and move the problem toward the common ground. Both the designer and the parties at stake continuously learn and reflect in a dialogical way. This dialogical learning is essential within participatory design, as participatory design sees people as the real experts of domains and experiences [74]. The notion of design of Simon [75] that design attempts to change existing situations into preferred ones transcends the designer's role in a complex setting; the whole network of stakeholders is necessary to get to the preferred situation. In a complex context, the preferable situation is inherently multileveled; therefore, it seems essential that different stakeholders reflect and expand their perspectives to formulate constructive objectives and inform future practice. A new *change space* might occur through dialogic reflection, where there is room for new ways of framing the problem by highlighting its paradoxes and eventually generating different possible solutions [76,77]. The reflective learning mechanism is often enacted by proposing or evaluating an intervention [3], which fits the nature of design by testing and assessing specific ideas, visualizations, concepts, and prototypes. The focus on social change and the emergence of a shared mental model regarding perspective making and perspective taking, informing future practice, might be a specific addition to the design process direction, providing social support to frame and reframe the problem. Unlike the identification mechanism, reflection is about overcoming boundaries and shaping future practice, where stakeholders are aware of the different perspectives resulting from perspective taking and perspective making.

Transformation

The dialogical learning mechanism transformation is about collaboration and the development or codevelopment of new practices [3]. The transformation mechanism is characterized by the process from a shared awareness of a problem to the development and, eventually, the crystallization of a new and maintainable setting. The ill-structured context becomes one where the innovation is characterized by tailored use. The emergence of a new context such as this is often the ultimate goal of both innovation and design. In the transformation phase, a shared problem space is necessary to get the whole network moving. Therefore, dialogical reflection in the system seems essential to advance a network to the transformation or change space.

Although there are design activities at the intersection of learning mechanisms, the learning mechanisms seem to be suitable for evaluating the degree of social change during the design process. In addition to the continuous reflection on the development of the product and the frame, it can be of added value to reflect on the social process early in the design process,

using learning mechanisms to increase the chances of integration and adoption.

Aims

Focusing on innovation in health, this study aims to find out whether the different learning mechanisms can be linked to studies on health innovation that mention boundary objects as a concept and assess whether the related mechanisms provide insight into the stage of the design and implementation or change process.

Methods

Databases and Search Strategy

The following six databases were searched for potentially relevant abstracts: PubMed, Scopus, Education Resources Information Center, PsycINFO, Information Science and Technology Abstracts, and Embase. These databases cover a wide range of published research in the field of health care. They were selected after several trial searches in various databases and after consultation with an information specialist in health science. The terms that were used for the search in PubMed are presented in Textbox 1.

Owing to differences in search engine functionality, the method by which terms were entered differed per database. A complete overview of the terms is included in Multimedia Appendix 1. Searches included papers published between 1989, when Star [1] introduced the concept of boundary objects, and September 2020. Before the definitive search, we performed 3 trial searches with different terms to reduce the possibility of missing relevant studies. We conducted a definitive search on September 23, 2020. We followed the PRISMA (Preferred Reporting Items of Systematic Reviews and Meta-Analyses) guidelines [78] as much as possible to report this review.

Textbox 1. Search terms used for relevant abstracts in PubMed.

Search terms

- (*boundary object*[tiab]* OR *boundary cross*[tiab]*) AND (*"Diffusion of Innovation" [Mesh]* OR *"Organizational Innovation" [Mesh]* OR *"Research" [Mesh]* OR *"Interdisciplinary Communication" [Mesh]* OR *"Negotiating" [Mesh]* OR *dialogic*[tiab]* OR *participatory[tiab]* OR *learn*[tiab]* OR *innovat*[tiab]* OR *design*[tiab]* OR *develop*[tiab]* OR *research*[tiab]* OR *interdisciplin*[tiab]* OR *cross disciplin*[tiab]* OR *multidisciplin*[tiab]* OR *negotiat*[tiab]* OR *mediat*[tiab]*)

Study Selection and Inclusion and Exclusion Criteria

We included studies that discussed boundary objects or innovations in health. We included only original reports or papers that (1) mentioned boundary objects, (2) involved an empirical study, or (3) otherwise focused on a newly developed or implemented innovation. Papers meeting these criteria were selected for full-text screening.

The following exclusion criteria were used for full-text screening: (1) non–peer-reviewed papers such as abstracts, conference posters, or trade journals; (2) papers with full text not available; (3) papers in languages other than English; (4) monographs or short reports; and (5) papers with not sufficient information in the abstract.

Screening Process

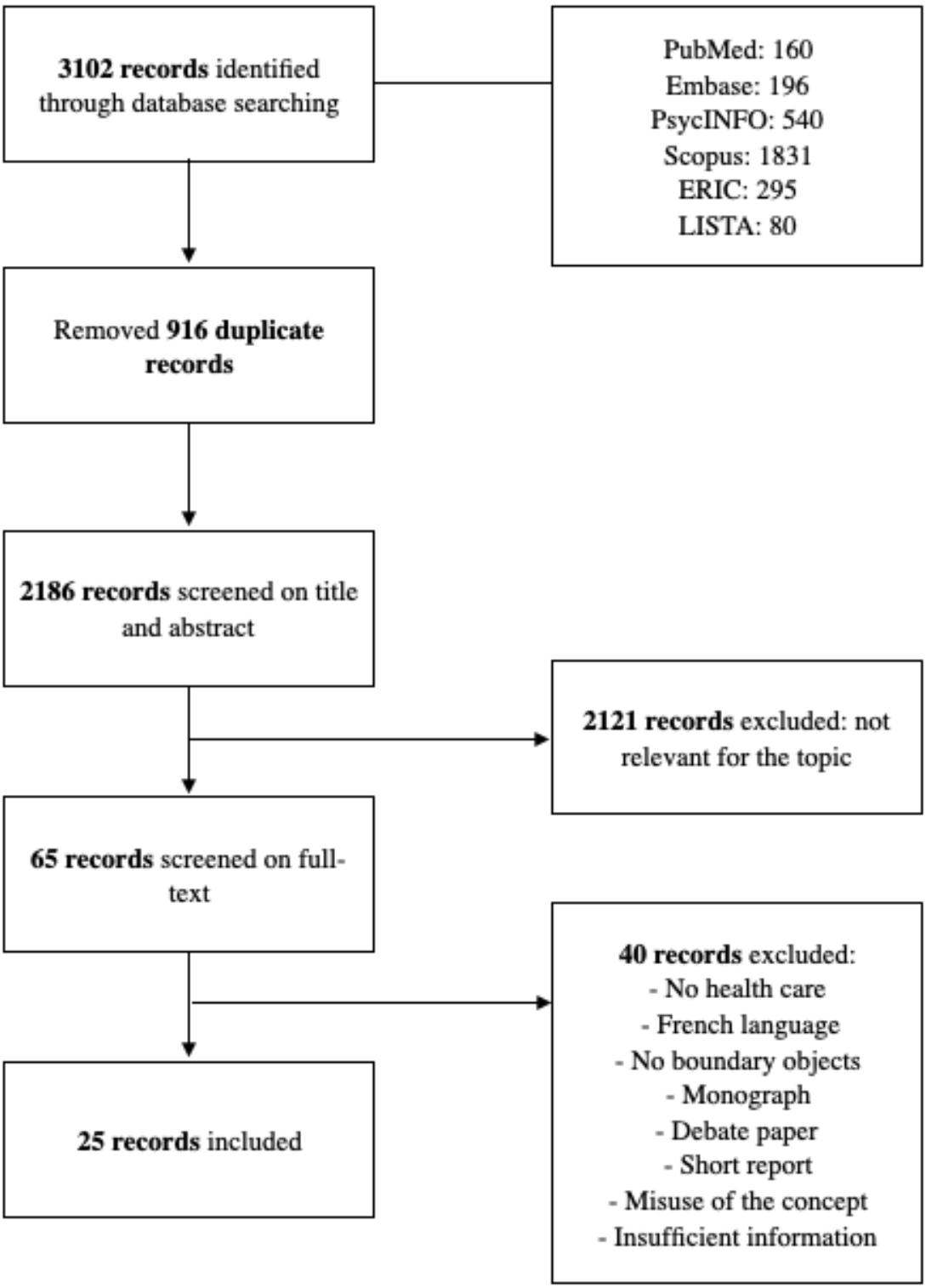
After removing the duplicates, the papers were screened based on title and abstract using Rayyan [79]. A total of 2 reviewers (GT and DK) independently reviewed all titles and abstracts, who were double-blinded for relevance with the formulated inclusion and exclusion criteria. Papers were only included on the agreement of both GT and DK, where a plausible argumentation for consideration of inclusion always led to inclusion. Full-text papers were retrieved after this step. During the full-text screening phase, the first 20% of the papers were randomly selected and double-blind reviewed by 2 reviewers (GT and LV). After this scan, no disagreements occurred about inclusion or the identified mechanisms. Then, the main reviewer (GT) reviewed the other included papers for a full-text reading.

Results

Search Results

Our initial search yielded 3102 records. Of the 3102 records, after removing the 916 (29.53%) duplicates, 2186 (70.47%) records were screened based on their titles and abstracts. Next, of the 65 records, we screened the full text, leaving 25 (38%) papers for inclusion (see Figure 1 for a flowchart of the results in the different selection stages). In both stages, there was a consensus between the reviewers on both the inclusion and analysis of the papers.

Figure 1. Flowchart of the selection process. ERIC: Education Resources Information Center; LISTA: Library, Information Science and Technology Abstracts.



General Findings

The studies included in this systematic review had varied study designs and topics. Table 1 presents the study designs, topics, and characteristics. All the included articles were published after 2008.

We studied full-text papers on applying the concept of boundary objects; whether this concept was used to describe daily life situations or situations where there was innovation either in its development, implementation, or postimplementation stage; and if ≥1 dialogical learning mechanisms could be identified. In the following section, we categorize the papers based on their mechanisms and the situations they applied to.

Table 1. Papers included in the systematic review.

Study	Title	Identified learning mechanisms	Phase
Nielsen and Mengiste, 2014 [80]	Analyzing the diffusion and adoption of mobile IT[a] across social worlds	• Identification • Transformation	Postimplementation
Lambert et al, 2019 [81]	Antimicrobial resistance, inflammatory responses: a comparative analysis of pathogenicities, knowledge hybrids and the semantics of antibiotic use	• Identification	Operational
Kajamaa, 2011 [82]	Boundary breaking in a hospital: expansive learning between the worlds of evaluation and frontline work	• Identification • Coordination • Reflection • Transformation	Development and implementation
Bjørn et al, 2009 [83]	Boundary factors and contextual contingencies: configuring electronic templates for health care professionals	• Identification	Development and implementation
Sajtos et al, 2018 [84]	Boundary objects for institutional work across service ecosystems	• Identification • Coordination • Reflection • Transformation	Implementation
Jensen and Kushniruk, 2016 [85]	Boundary objects in clinical simulation and design of eHealth	• Identification • Coordination • Reflection • Transformation	Development and implementation
Sampalli et al, 2011 [86]	Clinical vocabulary as a boundary object in multidisciplinary care management of multiple chemical sensitivity, a complex and chronic condition	• Coordination	Operational
Sampalli et al, 2009 [87]	Boundary objects in the multidisciplinary care management of chronic conditions: multiple chemical sensitivity	• Coordination	Operational
Fox, 2011 [88]	Boundary objects, social meanings and the success of new technologies	• Transformation	Postimplementation analysis
Jentoft, 2020 [89]	Boundary-crossings among health students in interprofessional geropsychiatric outpatient practice: collaboration with elderly people living at home	• Coordination • Reflection	Project
Håland et al, 2015 [90]	Care pathways as boundary objects between primary and secondary care: experiences from Norwegian home care services	• Coordination • Reflection • Transformation	Development and implementation
Sajtos et al, 2014 [91]	Case-mix system as a boundary object: the case of home care services	• Coordination	Development
Islind et al, 2019 [44]	Co-designing a digital platform with boundary objects: bringing together heterogeneous users in health care	• Identification • Coordination • Transformation	Development
Meier, 2015 [92]	Collaboration in health care through boundary work and boundary objects	• Coordination	Operational
Williams et al, 2008 [93]	Human embryos as boundary objects? Some reflections on the biomedical worlds of embryonic stem cells and preimplantation genetic diagnosis	• Identification	Operational
Keshet and Popper-Giveon, 2013 [94]	Integrative health care in Israel and traditional Arab herbal medicine: when health care interfaces with culture and politics	• Identification • Coordination • Reflection • Transformation	Operational
Marabelli et al, 2017 [95]	Knowledge sharing and health care coordination: the role of creation and use brokers	• Identification • Coordination	Development and implementation

Study	Title	Identified learning mechanisms	Phase
Islind and Snis, 2017 [96]	Learning in home care: a digital artifact as a designated boundary object-in-use	• Identification • Coordination • Transformation	Development and implementation
Isah and Bystroöm, 2020 [97]	The mediating role of documents: information sharing through medical records in health care	• Coordination	Operational
Gregory et al, 2014 [98]	Patient experiences of diabetes eHealth	• Coordination	Postimplementation
Stewart and Watson, 2019 [99]	A Sociotechnical history of the ultralightweight wheelchair: a vehicle of social change	• Transformation	Postimplementation
Melo and Bishop, 2020 [100]	Translating health care research evidence into practice: the role of linked boundary objects	• Coordination • Identification	Postimplementation and operational
Mengiste and Annestad, 2013 [101]	Understanding the dynamics of learning across social worlds: a case study from implementing IS[b] in the Ethiopian public health care system	• Identification • Coordination	Development
McLoughlin et al, 2016 [102]	Doing infrastructural work: the role of boundary objects in health information infrastructure projects	• Coordination	Development and implementation
Terlouw et al, 2020 [57]	Design of a digital comic creator (It's Me) to facilitate social skills training for children with autism spectrum disorder: design research approach	• Coordination • Reflection • Transformation	Development

[a]IT: information technology.

[b]IS: information system.

Identification

Of the 25 papers, 2 (8%) [81,93] described a medical term as a boundary object to identify different medical terms' interpretations over different contexts or disciplines. Lambert et al [81] approached terms such as *infection*, *antibiotics*, and *inflammation* as boundary objects. Williams et al [93] conceptualized embryos to find out how they were decontextualized and recontextualized within and between 2 different cultural systems. In both studies, existing beliefs were reinforced, and the studies aimed to identify local differences by applying the identification mechanism in a noninnovation context.

In 8% (2/25) of the papers, the identification mechanism was deduced in more innovative projects. Nielsen and Mengiste [80] described in their study an analysis of the adoption of mobile information technology innovation for home care. The innovation functioned as a boundary object on the level of influential stakeholders (ministry, local government, and managers). Managers and care workers revealed different interpretations of the technology's value and potential, resulting in resistance and tension. The technology seemed to reinforce existing differences between activity systems at the level of managers and care workers. The authors proposed a bottom-up approach and more involvement of end users in the future. This technology seemed to reinforce existing differences, triggering the identification mechanism.

Bjørn et al [83] described "conflicting perspectives between standardization and reconfiguration embedded within hospital information systems (HIS) design activities." In their study, the authors considered an electronic triage and tracking system as a boundary object. Users indicated whether they could work with the system and how. This led to adjustments and reconfigurations in the system and, presumably, better adoption of the system but not to, for example, perspective taking between different groups. Differences were primarily sought for user groups to use the system optimally by having them respond to the system, thus using the identification mechanism in this study in a constructive way to retrieve specific input.

Coordination

In 16% (4/25) of the papers, we identified the coordination mechanism to facilitate cooperation in daily practice. Approximately 50% (2/4) of these papers presented controlled clinical vocabulary to facilitate and coordinate collaboration between different professionals [86,87]. Approximately 50% (2/4) other studies make use of data or narratives to facilitate multidisciplinary cooperation. Meier [92] described an ethnographic study of 2 hospital wards. Patients' stories, especially their narratives and patient records, formed the boundary objects to make constructive multidisciplinary work possible. A study by Isah and Bystroöm [97] focused on the role of case notes as mediating artifacts in patient care. Their study demonstrated how case notes were a source of information and an essential, enacting, and mediating part of the work itself. The case notes seemed to be multipurpose; they served as a repository of information and knowledge and supported and mediated a plethora of the medical team's work activities in patient care. It was evident that the case notes served as a coordinating mechanism between the participating actors. Besides facilitating and fortifying many day-to-day functions in patient care, case notes have established themselves for

deliberate learning; they embody a clerkship template and enable newcomers to integrate and perpetuate the practice.

In 12% (3/25) of the papers, we identified a coordination mechanism to promote innovation. The coordination mechanism was identified in the first of 2 included papers from Sajtos et al [91], where the authors reported a case study of developing a case-mix system. Their study illustrated a process to address the diverse meanings and interests of various stakeholders to overcome communication and organizational challenges. They presented a funneling framework. A so-called boundary concept evolved through stakeholder input into a boundary object in a second step and a solution in the final step. Both principles and constraints were identified and addressed in the final solution by aligning the stakeholders' interests. In the project, no clients were directly involved, and it remains unclear whether and how the design itself was subject to flexibility along the way or that the project was mainly about fine-tuning a design.

Gregory et al [98] evaluated a diabetes eHealth system in their study. They described the use of the system as a boundary object for developing an understanding of why the eHealth system was used in a wide variety of ways, enabling coordination over stakeholder groups. McLoughlin et al [102] reported 4 case studies of health information infrastructure projects, where they approached the health information systems as boundary objects. Two of the projects were on a regional level, and 2 projects were on a national level. In the regional projects, the boundary objects managed and facilitated collaboration in using health data for different purposes by different users. At the national level, the 2 boundary objects instantiated top-down attempts and struggled more to trigger some effects.

Reflection

We identified no papers which merely reported the reflection mechanism.

Transformation

The transformation mechanism was identified in 8% (2/25) of the studies. Both studies described a more extended transformational development in retrospect. Fox [88] described the development of antiseptic and aseptic environments during surgery. In a historical case study, the report assessed the innovations in surgical sterility and how boundary objects worked over time. For example, in this study, nose masks were considered boundary objects in their relationship to social meanings within communities of practice. In the conclusions, the researcher described positive and negative boundary objects and concluded the following:

> *Boundary objects are not merely passive vehicles that allow communication between communities of practice or knowledge but elements that encapsulate the broader social meaning of a concept, theory, technology or practice, and the underlying relations surrounding its development and adoption.*

This study described the surgery profession's transformation process, partly through boundary objects, from only a healer of disease to a healer of disease who was also a safety procurer.

The second historical perspective was written by Stewart and Watson [99], who described the development of the ultralightweight wheelchair and its social implications. As a boundary object, the ultralightweight wheelchair had a significant transformational impact on the use of wheelchairs in the daily lives of users of wheelchairs. According to the authors, the wheelchair as a boundary object provided many insights through various interpretations of the artifact. It reflected views about users of wheelchairs and disability more generally and how the ultralightweight wheelchair as a boundary object seemed to manifest power relations between the diverse communities it engaged.

Multiple Mechanisms

Overview

Of the included 25 studies, in 12 (48%) studies, we identified ≥1 mechanism. Of these 12 papers, 3 (25%) focused more on structuring *everyday* practices, 9 (75%) identified multiple mechanisms that focused explicitly on developing or implementing a new tool, 6 (50%) were concentrated mainly on professionals and professional collaborations, and 3 (25%) actively focused on processes involving clients or patients at the center.

Studies Structuring Everyday Practice

Jentoft [89] described in her research boundary-crossing activities among physical therapy and medical students in interprofessional geropsychiatric outpatient practice. In the study, the students visited older clients living at home on 2 occasions. On the basis of these visits, the students considered suitable interventions for clients to enhance their quality of life, health outcomes, and well-being. After that, students wrote a health record to document their professional and interprofessional views on the cases. The health record and its content served as the boundary object during the study. The health record itself coordinated collaboration between the different disciplines by helping them plan examinations and establish a relationship with the client. The health record content enhanced reflection and negotiation and ensured that students understood the other's (professional) perspective better. In conclusion, the boundary objects led to more effectiveness and improved evaluation quality through better interprofessional collaboration, and students became more knowledgeable about what others and other professions did in practice.

Keshet and Popper-Giveon [94] explicitly used the learning mechanisms of Akkerman and Bakker [3] to describe the integration of local traditional medicine within complementary medicine. Their article aimed to contribute to the "contemporary critical debate in medical anthropology concerning medical pluralism and integrative medicine" by highlighting the exclusion of traditional medicine. Through ethnographic fieldwork, they focused on a group of integrative physicians who had recently begun integrating conventional herbal medicine. By conceptualizing traditional medicine as a boundary object, they attempted to bridge professional gaps between biomedicine, complementary medicine, and traditional medicine. Their study showed that using herbal medicine as a boundary

object helped overcome barriers and provided a window for dialog and learning at different levels.

Melo and Bishop [100] described a fall risk scale combined with a pink wristband as a boundary object. The pink wristband was used to signal patients with a high fall risk, measured by a falling score. Communicating the meaning of the pink wristband to other hospital staff improved the coordination and facilitation of work organization around persons with higher fall risk.

Studies Focusing Mainly on Professional Collaboration

This section describes the 24% (6/25) of papers where multiple identified mechanisms focused on developing or implementing a new tool to enhance or trigger professional collaboration. In 33% (2/6) of these papers, we discovered the identification and coordination mechanisms. Marabelli et al [95] described the development and implementation of a summary medical note (the single point of care]) carried by parents between the specialists involved in their child's care. Their paper described the single point of care as a boundary object with coordinative mechanics to enhance and facilitate communication between different stakeholders. In the predevelopment phase, parents had an important role in identifying and addressing the problem. The interviews and sessions had the characteristics to trigger the identification mechanism. In their analysis, the authors demonstrated that "the SPOC's effectiveness can be understood by looking at the combined roles of boundary objects and human brokers."

Mengiste and Annestad [101] reported a case study on the implementation of information systems in the Ethiopian public health care system. The paper analyzed how this software functioned as a boundary object. They found that the software did not just facilitate cooperation among the actors; the software as a boundary object also had a role in *bringing the existing differences to the foreground*, applying the identification and coordination mechanism. In addition to the software, which the authors explicitly called a boundary object, many sessions, workshops, test sessions, and prototypes were described, which also had the characteristics of boundary objects.

In the paper by Håland, Røsstad, and Osmundsen [90], coordinating and reflection mechanisms were identified. They studied the development, introduction, and use of "a care pathway across healthcare levels focusing on older home-dwelling patients in need of home care services after hospital discharge." Their study explored how care pathways can use the concept of boundary objects in translation between specialist health care services and home care services. Interviews with the project participants found that the "response to existing needs, local tailoring, involvement, and commitment are all crucial for the care pathway to function as a boundary object." Furthermore, they described that the artifact could "push boundaries just as much as it can be used as a tool for bridging across them" [90]. By introducing the care pathway system early, as an idea, to different stakeholders, they could address specific needs in the system, resulting in better integration. The introduction of the care pathway system led to collaboration and coordination among organizations, better understanding, reflection on different perspectives (eg, between home care workers and hospital care workers), and new ways of working in transformed activity systems.

In 12% (3/25) of the papers focusing on professionals and professional collaboration, we identified all 4 mechanisms. Jensen and Kushniruk [85] presented a case study on a participatory design process of electronic documentation templates for nurses, which they used for patient assessment:

> *Clinical simulation was used as a boundary object and thereby achieved mutual clinical agreement on the content. By using clinical simulation, knowledge was transferred and transformed between the different communities of practice to support gaining a shared understanding.*

This was mainly to overcome organizational barriers. As they presented in their case study, the clinical simulation might have helped form "shared mental models and shared understanding of user requirements, work practice and organizational requirements" within an innovation project. The boundary objects approach helped analyze vital issues and triggered a reflective approach to improving solutions. This case study showed that the adoption and acceptance of new technology might be significantly improved by leading end users and other important stakeholders within the organization through all mechanisms.

Sajtos et al [84] introduced the concept of boundary objects to facilitate institutional work across different ecosystems through a case-mix system. They conducted qualitative interviews with three key actors—funding agency, service provider, and clinicians—to identify these actors' views on the nature of home-based support services and their impact as a boundary object within the implementation of a case-mix system. Their analysis was based on three interviews: 1 before introducing the case-mix system, 1 just after the introduction, and 1 after the introduction. This provided a comprehensive view of an implementation process in which the concept of boundary objects was juxtaposed. The prephase mainly reported data reflecting the identification mechanism, where actors defined themselves mainly through differences between them. After the introduction, the case-mix system as a "boundary object enabled the actors to reframe and theorize about their idiosyncratic meanings of healthcare provision and embrace some new aspects." This led to perspective making and reframing of their own views to eventually use a jointly operated system by introducing new routines and practices that identified the reflection and transformation mechanisms. The reported study seemed to reflect the fluid implementation process by using the concept of boundary objects. The study did not report any adjustments made to the artifacts themselves because of the activated mechanisms or design rationale.

Kajamaa [82] reported a case study on the innovative creation process of an assessment tool in which nurses and quality controllers participated. Through different steps, the diverse needs of nurses and quality officers were reinforced and addressed. Both stakeholder groups collaborated on developing a tool, reflected on designs that led to perspective making and perspective taking, and finally started the implementation process together. The different *in-between* versions of the tool

acted as boundary objects. During implementation, 2 events occurred. The first event resulted from new circumstances, which were illustrative of solutions: problems are not static. This event was overcome during the project. The second event led to a breach of trust between the stakeholder groups and, thus, to the project's end. The initially overcome differences between the stakeholder groups were reinforced again by triggering the identification mechanism in a different way than the first time.

Studies Involving Clients or Patients

Of the 25 studies, 3 (12%) actively focused on processes involving clients or patients at the center. In these studies, clients or patients actively participated, and 3 mechanisms were identified. Islind et al [44] applied the concept of boundary objects in a co-design project for a digital platform at a clinic that supported cancer patients in their struggles with treatment-induced illnesses. This paper explicitly explored the functions that boundary objects can have in a design process and how they were engaged in the different design phases. Islind et al [44] described the following three types of boundary objects: narratives as open boundary objects in the first phase, metaphorical boundary objects as semiopen boundary objects in the second phase, and structured boundary objects in the third phase. Although the focus was more on the boundary objects' different characteristics during a design project, implicitly, the mechanisms that the boundary objects enacted were also described. The first type of boundary objects—the narratives—seemed to trigger the identification mechanism to better understand the user groups:

> *The narratives, in the forms of patient stories, played a central role for understanding the patient group and the healthcare professionals as the needs of both user groups needed to be accommodated for.*
>
> *In a way, the narrative became the container of the essence of being a patient.*

In what Islind et al [44] called the *metaphorical phase*, boundary objects facilitated conversation, collaboration, and consultation among stakeholders, aligning with the coordination mechanism. In the structured phase, the boundary objects matured more as prototypes. They triggered a conversation about the platform's future functions, aligning with the first signs of transformation. Their conclusion stated the following:

> *Designing with boundary objects might slow down the design process initially but actually speed up the programming process as fewer aspects will come as a surprise during the software development when everything has been negotiated thoroughly on beforehand.*

In a study by Islind and Snis [96], the focus was on developing and deploying a mobile health (mHealth) artifact for groceries in home care settings. An mHealth artifact "was tested to see how the quality of home care work practice was enhanced and changed." The mHealth artifact was presented in this paper as a boundary object. The authors presented the artifact as a designated boundary object and a boundary object in use. As a boundary object, the mHealth artifact triggered different mechanisms. In conversations, the tool reinforced the identity of older adults. For example, they realized how long they had not been to a grocery store. From the older adults' perspective, the boundary object functioned as "a substitute for their previous buying groceries." From the caregivers' perspective, the boundary object was designed to "support a more efficient working process," triggering both coordinate and transformational mechanisms. The time earlier spent in the grocery store now went to the older adults, leading to more caregiving quality in praxis. The mHealth tool was described as follows:

> *Mediating tool for a deepened caring conversation-in-practice where interactions and realizations generate new emerging properties and opportunities. The boundary object-in-use proved to function as a conversation starter where the use facilitated fruitful conversations between the elderly and caregivers about new aspects of grocery shopping.*

In addition, new diet and "nutrition explorations were interpreted and negotiated via their evolved conversation." This reshaping of the home care practice affected the caregivers' role, "evolving into a more meaningful caretaking and nurturing role."

Terlouw et al [57] described the development of a digital comic creator for children with an autism spectrum disorder. The digital tool was approached during the process and designed as a boundary object, aiming to connect the different stakeholders' objectives. This led to an inclusive design and triggered reflection and transformation learning mechanisms along the way.

Discussion

Boundary Objects in Health

This review shows that the concept of boundary objects has found its way into health care. The use of the concept has been growing since 2008, with a significant number of papers describing boundary objects from the past 5 years. In the reviewed studies, we see that boundary objects are mainly used to shape and organize multidisciplinary work, close to the original explanation of Star and Griesemer [46], or to surface differences in, for example, interpretation of a concept from different contexts or disciplines. In the 25 papers, 38 mechanisms were identified, of which 15 (39%) were coordination mechanisms, and 10 (26%) were identification mechanisms. In addition to the organizing and performative effect, boundary objects can reinforce boundaries and create conflicts. In addition to the proposition by Star and Griesemer [46], Oswick and Robertson [103] referred to *barricades and mazes* that generate conflict and reinforce boundaries and existing differences, something that Langley et al [48] also described as part of competitive boundary work. This can be an opposing and perhaps unwelcome side of the identification mechanism in terms of change management. In the study by Kajamaa [82], we saw this effect. First, in what seems a fluid development and implementation process, they applied the identification mechanism to identify different stakeholders' needs. After implementation, 1 event led to a breakdown of

trust between stakeholders, which led to the project's withdrawal. After this event, the boundary object was primarily used to name the significant differences between stakeholders and compete for a position without the other.

Although the concept of a boundary object was introduced to describe how specific artifacts can fulfill a bridging function between different sociocultural sites [1] and thus, have a social focus, the focus in the included papers was often on the boundary object itself rather than the social effect. Various labels were given to boundary objects in different studies, which described a more designerly process for an artifact. In the different included and excluded studies, we saw a differentiation between *designated boundary objects* and *boundary objects in use* [96,104,105]. This differentiation can be seen as parallel to the design research process. An artifact or solution continues to take shape and is developed in small steps from prototype to *object in use*. The analogy can also be made by applying a boundary object from a more ill-structured to a more well-structured context. In a second included paper of Islind et al [44], they described three types of boundary objects: narratives as open boundary objects in the first phase, metaphorical boundary objects as semiopen boundary objects in the second phase, and structured boundary objects in the third phase. Although again, the focus was more on the development of the object itself rather than the effects of the object in the social context, parallels can be drawn with a design process and application of the learning mechanisms in practice, as can be seen in the results.

The reflection and transformation mechanisms are underrepresented in the included studies. Of the 25 studies, 2 (8%) describe the transformative effect of boundary objects from a historical perspective [88,99], describing a long timeline of a particular development. However, it is difficult to determine the impact of the boundary object itself in retrospect as it is likely that many more variables played a role in the transformational processes. In addition, it is difficult to determine, in retrospect, whether the boundary objects were deliberately deployed for the given purpose. The reflective mechanism was the least identified in all the papers. However, in the papers in which the reflective mechanism took place [57,82,84,85,90], there was a much smoother adaptation and application of the innovation or tool afterward. There was more shared ownership of the problem and solution in the processes described and more consideration of other perspectives along the way. This reinforces the idea of reflection as an essential step in the design process, especially in a more complex setting with multiple stakeholders, needs, and interests. When these are appropriately addressed in the design through a boundary object's focus and, simultaneously, addressed within the design, more mutual understanding arises. This leads to a natural emerging change space where everyone is willing to move forward [76,77].

On the basis of the findings of this study, for future design and implementation projects, the social focus of boundary objects can add value to innovation projects. Pursuing to trigger the reflective mechanism can lead to the benefit of more fluid and smooth integration of innovation into practice. Here, the boundary object perspective avoids the pursuit of consensus, which often proves unfeasible in complex practices with many stakeholders. The reflective mechanism creates a shared awareness that there are multiple perspectives and needs. This awareness can lead to a shared change space in which innovation can flourish.

Strengths and Limitations

As seen in previous research [106], little attention has been paid to describing a conscious rationale for designing innovative artifacts in health care research. This makes it hard to determine the thoughts and foundations of a designed object. In this study, this fact also made it difficult to ascertain the intent behind the deployment of particular boundary objects. The effect was often identifiable; however, it was impossible to determine whether it was directed or accidental without knowing the intention. In addition, no study described what changes were explicitly made to a prototype or design after a specific stakeholder workshop or meeting. The often implicit focus on effect is evident in health care research, making it difficult for innovative design processes to get sensible insight into the design rationales of others.

Another observation was that many innovations in the included studies were more administrative systems, such as electronic patient files. These are pre-eminently systems with which different disciplines must work, and boundary objects are thus helpful; however, 12% (3/25) of studies showed that boundary objects are also of added value in research in which clients or patients have an active role. This observation raises the idea that there are still more gains that can be found by involving end users earlier in design processes.

The included papers were subject to the interpretation, discussion, and consensus of the reviewers (GT, DK, and LV). To counteract subjectivity as much as possible, papers were double-blind reviewed by 2 reviewers in the title and abstract scan (GT and DK). They were only included in the consensus of both reviewers. In the full-read phase, 20% (13/65) of the papers were double-blind reviewed by 2 reviewers (GT and LV) before they were discussed. No disagreements on inclusion occurred during the discussion.

Conclusions

The concept of boundary objects has found its way into health care. In this review, we saw that boundary objects in health are primarily used to shape and organize multidisciplinary work or to surface differences in, for example, the interpretation of a concept from different contexts or disciplines. Although the concept of a boundary object was introduced to describe how specific artifacts can fulfill a bridging function between different sociocultural sites and thus have a social focus, the focus in the included papers was often on the boundary object itself rather than the social effect. The reflection and transformation mechanisms were underrepresented in the included studies; however, based on the findings in this review, pursuing to trigger the reflective mechanism in design, development, and implementation projects can lead to the benefit of more fluid and smooth integration of innovation into practice.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search keys used per database.
[DOCX File , 13 KB - humanfactors_v9i1e31167_app1.docx]

References

1. Star S. The Structure of Ill-Structured Solutions: boundary objects and heterogeneous distributed problem solving. Distrib Artif Intell 1989:37-54. [doi: 10.1016/b978-1-55860-092-8.50006-x]
2. Engeström Y, Engeström R, Kärkkäinen M. Polycontextuality and boundary crossing in expert cognition: learning and problem solving in complex work activities. Learn ing and Instruct 1995 Jan;5(4):319-336. [doi: 10.1016/0959-4752(95)00021-6]
3. Akkerman SF, Bakker A. Boundary crossing and boundary objects. Rev Edu Res 2011 Jun 01;81(2):132-169 [FREE Full text] [doi: 10.3102/0034654311404435]
4. Carlile PR. A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development. Organization Science 2002 Aug;13(4):442-455 [FREE Full text] [doi: 10.1287/orsc.13.4.442.2953]
5. Carlile PR. Transferring, translating, and transforming: an integrative framework for managing knowledge across boundaries. Organ Sci 2004 Oct;15(5):555-568. [doi: 10.1287/orsc.1040.0094]
6. Nicolini D, Mengis J, Swan J. Understanding the role of objects in cross-disciplinary collaboration. Organ Sci 2012 Jun;23(3):612-629. [doi: 10.1287/orsc.1110.0664]
7. Leigh Star S. This is not a boundary object: reflections on the origin of a concept. Sci Technol Hum Val 2010 Aug 10;35(5):601-617. [doi: 10.1177/0162243910377624]
8. Alexander J, Hearld L. The science of quality improvement implementation: developing capacity to make a difference. Med Care 2011 Dec;49 Suppl:6-20. [doi: 10.1097/MLR.0b013e3181e1709c] [Medline: 20829724]
9. Jacobs SR, Weiner BJ, Reeve BB, Hofmann DA, Christian M, Weinberger M. Determining the predictors of innovation implementation in healthcare: a quantitative analysis of implementation effectiveness. BMC Health Serv Res 2015 Jan 22;15(1):6 [FREE Full text] [doi: 10.1186/s12913-014-0657-3] [Medline: 25608564]
10. Kellermann AL, Jones SS. What it will take to achieve the as-yet-unfulfilled promises of health information technology. Health Aff (Millwood) 2013 Jan;32(1):63-68. [doi: 10.1377/hlthaff.2012.0693] [Medline: 23297272]
11. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471. [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
12. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011 Jan 18;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
13. Goldzweig CL, Towfigh A, Maglione M, Shekelle PG. Costs and benefits of health information technology: new trends from the literature. Health Aff (Millwood) 2009 Jan;28(2):282-293. [doi: 10.1377/hlthaff.28.2.w282] [Medline: 19174390]
14. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
15. Rogers EM. Diffusion of Innovations, 5th Edition. New York: Simon & Schuster; 2003.
16. Lavis JN, Robertson D, Woodside JM, McLeod CB, Abelson J, Knowledge Transfer Study Group. How can research organizations more effectively transfer research knowledge to decision makers? Milbank Q 2003 Jun 06;81(2):171-2; 221-48 [FREE Full text] [doi: 10.1111/1468-0009.t01-1-00052] [Medline: 12841049]
17. Klein KJ, Sorra JS. The challenge of innovation implementation. Acad Manag Rev 1996 Oct;21(4):1055-1080. [doi: 10.5465/amr.1996.9704071863]
18. Blase K, Fixsen D. The national implementation research network: improving the science and practice of implementation. PsycEXTRA Dataset 2005. [doi: 10.1037/e522472006-003]
19. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A, "Psychological Theory" Group. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb 01;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
20. Lomas J. Diffusion, dissemination, and implementation: who should do what? Ann N Y Acad Sci 1993 Dec 31;703(1 Doing More Go):226-235. [doi: 10.1111/j.1749-6632.1993.tb26351.x] [Medline: 8192299]
21. Graham I, Logan J, Harrison M, Straus S, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]

22. Moullin JC, Sabater-Hernández D, Fernandez-Llimos F, Benrimoj SI. A systematic review of implementation frameworks of innovations in healthcare and resulting generic implementation framework. Health Res Policy Syst 2015 Mar 14;13(1):16 [FREE Full text] [doi: 10.1186/s12961-015-0005-z] [Medline: 25885055]
23. Manzini E, Coad R. Design, When Everybody Designs: An Introduction to Design for Social Innovation. Cambridge, MA: MIT Press; 2015.
24. Brown T. Change by Design, Revised and Updated: How Design Thinking Transforms Organizations and Inspires Innovation. New York: Harper Business; 2019.
25. Dorst K. Frame Innovation: Create New Thinking by Design. Cambridge, MA: MIT Press; 2015.
26. Partridge R. Understanding the roles of the designer in health care: a practice-based study into supporting adolescents with long-term conditions. Design J 2017 Jun 13;20(4):523-532. [doi: 10.1080/14606925.2017.1325586]
27. Tsekleves E, Cooper R. Emerging trends and the way forward in design in healthcare: an expert's perspective. Design J 2017 Sep 06;20(sup1):2258-2272. [doi: 10.1080/14606925.2017.1352742]
28. Wildevuur SE. Could health learn from design? Design Health 2017 Mar 13;1(1):59-64. [doi: 10.1080/24735132.2017.1295707]
29. Ekman I, Swedberg K, Taft C, Lindseth A, Norberg A, Brink E, et al. Person-centered care--ready for prime time. Eur J Cardiovasc Nurs 2011 Dec;10(4):248-251. [doi: 10.1016/j.ejcnurse.2011.06.008] [Medline: 21764386]
30. Pollard S, Bansback N, Bryan S. Physician attitudes toward shared decision making: a systematic review. Patient Educ Couns 2015 Sep;98(9):1046-1057. [doi: 10.1016/j.pec.2015.05.004] [Medline: 26138158]
31. Hassenzahl M, Diefenbach S, Göritz A. Needs, affect, and interactive products – Facets of user experience. Interact Comput 2010 Sep;22(5):353-362. [doi: 10.1016/j.intcom.2010.04.002]
32. Friedman B. Value-sensitive design. Interactions 1996 Dec;3(6):16-23. [doi: 10.1145/242485.242493]
33. Vines J, Clarke R, Wright P, McCarthy J, Olivier P. Configuring participation: on how we involve people in design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; April 27- May 2, 2013; Paris France. [doi: 10.1145/2470654.2470716]
34. Østergaard KL, Simonsen J, Karasti H. Examining situated design practices: nurses' transformations towards genuine participation. Design Stud 2018 Nov;59:37-57. [doi: 10.1016/j.destud.2017.12.002]
35. Bate P, Robert G. Experience-based design: from redesigning the system around the patient to co-designing services with the patient. Qual Saf Health Care 2006 Oct 01;15(5):307-310 [FREE Full text] [doi: 10.1136/qshc.2005.016527] [Medline: 17074863]
36. Sanders L, Stappers PJ. Convivial Toolbox: Generative Research for the Front End of Design. The Netherlands: BIS Publishers; 2012.
37. The field guide to human-centered design. Design Kit. URL: https://www.designkit.org/resources/1 [accessed 2022-01-25]
38. Stickdorn M, Schneider J. This is Service Design Thinking: Basics, Tools, Cases. The Netherlands: BIS Publishing; 2010.
39. Lin M, Hughes B, Katica M, Dining-Zuber C, Plsek P. Service design and change of systems: human-centered approaches to implementing and spreading service design. Int J Design 2011;5(2):73-86 [FREE Full text]
40. Curedale R. Design Thinking Process and Methods 5th Edition. California (US): Design Community College Inc; 2019.
41. Kuipers DA, Terlouw G, Wartena BO, Prins JT, Pierie JP. Maximizing authentic learning and real-world problem-solving in health curricula through psychological fidelity in a game-like intervention: development, feasibility, and pilot studies. Med Sci Educ 2019 Mar 12;29(1):205-214 [FREE Full text] [doi: 10.1007/s40670-018-00670-5] [Medline: 34457469]
42. Kuipers DA, Terlouw G, Wartena BO, van Dongelen R. Design for vital regions lessons learned uit Digitale Innovatie in Zorg and Welzijn. In: Coenders M, Metselaar J, Thijssen J, editors. Vital Regions: Samen Bundelen van Praktijkgericht Onderzoek. Leeuwarden: NHL Stenden Hogeschool; 2018:251-261.
43. Stickdorn M, Hormess ME, Lawrence A, Schneider J. This is Service Design Doing: Applying Service Design Thinking in the Real World. Newton, Massachusetts, United States: O'Reilly Media Inc; 2018.
44. Islind AS, Lindroth T, Lundin J, Steineck G. Co-designing a digital platform with boundary objects: bringing together heterogeneous users in healthcare. Health Technol 2019 May 20;9(4):425-438. [doi: 10.1007/s12553-019-00332-5]
45. Rhinow H, Köppen E, Meinel C. Design prototypes as boundary objects in innovation processes. In: Proceedings of the International Conference on Design Research Society. 2012 Presented at: International Conference on Design Research Society; July 1-4, 2012; Bangkok, Thailand URL: https://dl.designresearchsociety.org/drs-conference-papers/drs2012/researchpapers/116/
46. Star SL, Griesemer JR. Institutional ecology, `Translations' and boundary objects: amateurs and professionals in Berkeley's museum of vertebrate zoology, 1907-39. Soc Stud Sci 2016 Jun 29;19(3):387-420 [FREE Full text] [doi: 10.1177/030631289019003001]
47. Dorst K. Design beyond design. She Ji J Design Econ Innov 2019;5(2):117-127. [doi: 10.1016/j.sheji.2019.05.001]
48. Langley A, Lindberg K, Mørk BE, Nicolini D, Raviola E, Walter L. Boundary work among groups, occupations, and organizations: from cartography to process. Ann R Coll Physicians Surg Can 2019 Jul;13(2):704-736. [doi: 10.5465/annals.2017.0089]

49. Brown T. Design Thinking. 2008 Jun. URL: https://readings.design/PDF/Tim%20Brown,%20Design%20Thinking.pdf [accessed 2021-09-09]
50. Kolko J. Exposing the Magic of Design: A Practitioner's Guide to the Methods and Theory of Synthesis. Oxford, United Kingdom: Oxford University Press; 2011.
51. Jain J. Design Thinking for Startups: A Handbook for Readers and Workbook for Practitioners. California: Notion Press; 2018.
52. 11 lessons: managing design in 11 global brands : a study of the design process. Design Council. 2007. URL: https://www.designcouncil.org.uk/resources/report/11-lessons-managing-design-global-brands [accessed 2020-08-20]
53. Hardy A, Wojdecka A, West J, Matthews E, Golby C, Ward T, et al. How inclusive, user-centered design research can improve psychological therapies for psychosis: development of SlowMo. JMIR Ment Health 2018 Dec 05;5(4):e11222 [FREE Full text] [doi: 10.2196/11222] [Medline: 30518514]
54. Clune SJ, Lockrey S. Developing environmental sustainability strategies, the Double Diamond method of LCA and design thinking: a case study from aged care. J Clean Prod 2014 Dec;85:67-82. [doi: 10.1016/j.jclepro.2014.02.003]
55. Kuipers DA, Wartena BO, Dijkstra BH, Terlouw G, van T Veer JT, van Dijk HW, et al. iLift: a health behavior change support system for lifting and transfer techniques to prevent lower-back injuries in healthcare. Int J Med Inform 2016 Dec;96:11-23. [doi: 10.1016/j.ijmedinf.2015.12.006] [Medline: 26797571]
56. Terlouw G, van't Veer JT, Kuipers DA, Metselaar J. Context analysis, needs assessment and persona development: towards a digital game-like intervention for high functioning children with ASD to train social skills. Early Child Dev Care 2018 Dec 15;190(13):2050-2065. [doi: 10.1080/03004430.2018.1555826]
57. Terlouw G, van 't Veer JT, Prins JT, Kuipers DA, Pierie JE. Design of a digital comic creator (It's Me) to facilitate social skills training for children with autism spectrum disorder: design research approach. JMIR Ment Health 2020 Jul 10;7(7):e17260 [FREE Full text] [doi: 10.2196/17260] [Medline: 32673273]
58. Rogers CR. Empathic: an unappreciated way of being. Counsel Psychol 2016 Sep 04;5(2):2-10. [doi: 10.1177/001100007500500202]
59. Gray D, Brown S, Macanufo J. Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers. Sebastopol, California: O'Reilly Media, Inc; 2010.
60. Nilsson AW, Fältholm Y, Abrahamsson L. Reframing practice through the use of Personas. Reflect Pract 2010 Jun 18;11(3):285-298. [doi: 10.1080/14623943.2010.487376]
61. So C, Joo J. Does a Persona improve creativity? Design J 2017 May 14;20(4):459-475. [doi: 10.1080/14606925.2017.1319672]
62. Kahneman D, Krueger AB, Schkade DA, Schwarz N, Stone AA. A survey method for characterizing daily life experience: the day reconstruction method. Science 2004 Dec 03;306(5702):1776-1780. [doi: 10.1126/science.1103572] [Medline: 15576620]
63. van der Bijl-Brouwer M, Dorst K. Advancing the strategic impact of human-centred design. Design Stud 2017 Nov;53:1-23. [doi: 10.1016/j.destud.2017.06.003]
64. Veldmeijer L, Wartena B, Terlouw G, van't Veer J. Reframing loneliness through the design of a virtual reality reminiscence artefact for older adults. Design Health 2020 Nov 26;4(3):407-426. [doi: 10.1080/24735132.2020.1848976]
65. Poncette A, Rojas P, Hofferbert J, Sosa AV, Balzer F, Braune K. Hackathons as stepping stones in health care innovation: case study with systematic recommendations. J Med Internet Res 2020 Mar 24;22(3):e17004 [FREE Full text] [doi: 10.2196/17004] [Medline: 32207691]
66. Walker A, Ko N. Bringing medicine to the digital age via hackathons and beyond. J Med Syst 2016 Apr 12;40(4):98. [doi: 10.1007/s10916-016-0461-1] [Medline: 26872780]
67. Pathanasethpong A, Soomlek C, Morley K, Morley M, Polpinit P, Dagan A, et al. Tackling regional public health issues using mobile health technology: event report of an mHealth hackathon in Thailand. JMIR Mhealth Uhealth 2017 Oct 16;5(10):e155 [FREE Full text] [doi: 10.2196/mhealth.8259] [Medline: 29038098]
68. Vavoula GN, Sharples M. Future technology workshop: a collaborative method for the design of new learning technologies and activities. Comput Supp Learn 2007 Oct 16;2(4):393-419. [doi: 10.1007/s11412-007-9026-0]
69. Lupton D. Digital health now and in the future: findings from a participatory design stakeholder workshop. Digit Health 2017 Nov 09;3:2055207617740018 [FREE Full text] [doi: 10.1177/2055207617740018] [Medline: 29942616]
70. Noergaard B, Sandvei M, Rottmann N, Johannessen H, Wiil U, Schmidt T, et al. Development of a web-based health care intervention for patients with heart disease: lessons learned from a participatory design study. JMIR Res Protoc 2017 May 17;6(5):e75 [FREE Full text] [doi: 10.2196/resprot.7084] [Medline: 28526674]
71. Pfisterer KJ, Boger J, Wong A. Prototyping the automated food imaging and nutrient intake tracking system: modified participatory iterative design sprint. JMIR Hum Factors 2019 May 09;6(2):e13017 [FREE Full text] [doi: 10.2196/13017] [Medline: 31094336]
72. Moore G, Wilding H, Gray K, Castle D. Participatory methods to engage health service users in the development of electronic health resources: systematic review. J Particip Med 2019 Feb 22;11(1):e11474 [FREE Full text] [doi: 10.2196/11474] [Medline: 33055069]
73. Schön DA. The Reflective Practitioner: How Professionals Think In Action. New York: Routledge; 2017.

74. Sanders L. ON MODELING: an evolving map of design practice and design research. Interactions 2008 Nov 01;15(6):13-17. [doi: 10.1145/1409040.1409043]
75. Simon HA. Sciences of the Artificial. Cambridge, MA: MIT Press; 1969.
76. Dorst K. The core of 'design thinking' and its application. Design Stud 2011 Nov;32(6):521-532. [doi: 10.1016/j.destud.2011.07.006]
77. Dorst K. Frame creation and design in the expanded field. She Ji J Design Econ Innov 2015;1(1):22-33. [doi: 10.1016/j.sheji.2015.07.003]
78. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. Br Med J 2009 Jul 21;339(jul21 1):b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
79. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
80. Nielsen JA, Mengiste SA. Analysing the diffusion and adoption of mobile IT across social worlds. Health Informatics J 2014 Jun 08;20(2):87-103 [FREE Full text] [doi: 10.1177/1460458213481688] [Medline: 24810724]
81. Lambert H, Chen M, Cabral C. Antimicrobial resistance, inflammatory responses: a comparative analysis of pathogenicities, knowledge hybrids and the semantics of antibiotic use. Palgrave Commun 2019 Jul 30;5(1). [doi: 10.1057/s41599-019-0293-y]
82. Kajamaa A. Boundary breaking in a hospital. Learn Organ 2011 Jul 19;18(5):361-377. [doi: 10.1108/09696471111151710]
83. Bjørn P, Burgoyne S, Crompton V, MacDonald T, Pickering B, Munro S. Boundary factors and contextual contingencies: configuring electronic templates for healthcare professionals. Eur J Inform Sys 2017 Dec 19;18(5):428-441. [doi: 10.1057/ejis.2009.34]
84. Sajtos L, Kleinaltenkamp M, Harrison J. Boundary objects for institutional work across service ecosystems. J Serv Manag 2018 Jul 13;29(4):615-640. [doi: 10.1108/josm-01-2017-0011]
85. Jensen S, Kushniruk A. Boundary objects in clinical simulation and design of eHealth. Health Informatics J 2016 Jun 09;22(2):248-264 [FREE Full text] [doi: 10.1177/1460458214551846] [Medline: 25301197]
86. Sampalli T, Shepherd M, Duffy J. Clinical vocabulary as a boundary object in multidisciplinary care management of multiple chemical sensitivity, a complex and chronic condition. J Multi Discip Healthc 2011 Apr:91. [doi: 10.2147/jmdh.s17564]
87. Sampalli T, Shepherd M, Fox R. Boundary objects in the multidisciplinary care management of chronic conditions: multiple chemical sensitivity. Stud Health Technol Inform 2009;143:534-539. [Medline: 19380988]
88. Fox NJ. Boundary objects, social meanings and the success of new technologies. Sociology 2011 Feb 04;45(1):70-85. [doi: 10.1177/0038038510387196]
89. Jentoft R. Boundary-crossings among health students in interprofessional geropsychiatric outpatient practice: collaboration with elderly people living at home. J Interprof Care 2021 Apr 01;35(3):409-418. [doi: 10.1080/13561820.2020.1733501] [Medline: 32233885]
90. Håland E, Røsstad T, Osmundsen TC. Care pathways as boundary objects between primary and secondary care: experiences from Norwegian home care services. Health (London) 2015 Nov 26;19(6):635-651. [doi: 10.1177/1363459314567789] [Medline: 25627695]
91. Sajtos L, Rouse P, Harrison J, Parsons M. Case-mix system as a boundary object: the case of home care services. Australas Mark J 2021 Jan 27;22(3):189-196. [doi: 10.1016/j.ausmj.2014.08.003]
92. Meier N. Collaboration in healthcare through boundary work and boundary objects. Qual Sociol Rev 2015 Jul 31;11(3):60-82 [FREE Full text] [doi: 10.18778/1733-8077.11.3.05]
93. Williams C, Wainwright SP, Ehrich K, Michael M. Human embryos as boundary objects? Some reflections on the biomedical worlds of embryonic stem cells and pre-implantation genetic diagnosis. New Genet Society 2008 May 07;27(1):7-18. [doi: 10.1080/14636770701843550]
94. Keshet Y, Popper-Giveon A. Integrative health care in Israel and traditional arab herbal medicine: when health care interfaces with culture and politics. Med Anthropol Q 2013 Sep 08;27(3):368-384. [doi: 10.1111/maq.12049] [Medline: 24248993]
95. Marabelli M, Newell S, Krantz C, Swan J. Knowledge sharing and health-care coordination: the role of creation and use brokers. Health Systems 2017 Dec 19;3(3):185-198. [doi: 10.1057/hs.2014.8]
96. Islind AS, Lundh Snis U. Learning in home care: a digital artifact as a designated boundary object-in-use. J Workplace Learn 2017 Sep 11;29(7/8):577-587. [doi: 10.1108/jwl-04-2016-0027]
97. Isah EE, Byström K. The mediating role of documents: information sharing through medical records in healthcare. J Document 2020 Apr 04;76(6):1171-1191. [doi: 10.1108/jd-11-2019-0227]
98. Gregory P, Byrne P, Gabbay M. Patient experiences of diabetes eHealth. Int J Sociotechnol Knowled Dev (IJSKD) 2014;6:17. [doi: 10.4018/ijskd.2014010101]
99. Stewart H, Watson N. A sociotechnical history of the ultralightweight wheelchair: a vehicle of social change. Sci Technol Hum Val 2019 Dec 06;45(6):1195-1219. [doi: 10.1177/0162243919892558]
100. Melo S, Bishop S. Translating healthcare research evidence into practice: the role of linked boundary objects. Soc Sci Med 2020 Feb;246:112731. [doi: 10.1016/j.socscimed.2019.112731] [Medline: 31864175]
101. Mengiste SA, Aanestad M. Understanding the dynamics of learning across social worlds: a case study from implementing IS in the Ethiopian public health care system. Inform Organ 2013 Oct;23(4):233-257. [doi: 10.1016/j.infoandorg.2013.08.001]

102. McLoughlin I, Garrety K, Wilson R, Dalley A, Yu P. Doing infrastructural work: the role of boundary objects in health information infrastructure projects. Scandinavian J Inform Sys 2016;34(2):1-29 [FREE Full text]
103. Oswick C, Robertson M. Boundary objects reconsidered: from bridges and anchors to barricades and mazes. J Change Manag 2009 Jun;9(2):179-193. [doi: 10.1080/14697010902879137]
104. Garrety K, Dalley A, McLoughlin I, Wilson R, Yu P. Managing collaboration across boundaries in health information technology projects. Stud Health Technol Inform 2012;178:58-63. [Medline: 22797020]
105. Levina, Vaast. The emergence of boundary spanning competence in practice: implications for implementation and use of information systems. MIS Q 2005;29(2):335. [doi: 10.2307/25148682]
106. Kuipers DA, Terlouw G, Wartena BO, van 't Veer JT, Prins JT, Pierie JP. The role of transfer in designing games and simulations for health: systematic review. JMIR Serious Games 2017 Nov 24;5(4):e23 [FREE Full text] [doi: 10.2196/games.7880] [Medline: 29175812]

Abbreviations

mHealth: mobile health
PRISMA: Preferred Reporting Items of Systematic Reviews and Meta-Analyses

Edited by A Kushniruk; submitted 11.06.21; peer-reviewed by A Islind, S Hewko; comments to author 26.08.21; revised version received 15.10.21; accepted 17.10.21; published 03.02.22.

Please cite as:
Terlouw G, Kuipers D, Veldmeijer L, van 't Veer J, Prins J, Pierie JP
Boundary Objects as Dialogical Learning Accelerators for Social Change in Design for Health: Systematic Review
JMIR Hum Factors 2022;9(1):e31167
URL: https://humanfactors.jmir.org/2022/1/e31167
doi:10.2196/31167
PMID:35113023

Review

Status of Compassionate, Respectful, and Caring Health Service Delivery: Scoping Review

Adane Nigusie[1*], MPH; Berhanu F Endehabtu[2], MPH; Dessie Abebaw Angaw[3], MPH; Alemayehu Teklu[4], MD; Zeleke Abebaw Mekonnen[2], MPH; Marta Feletto[5], PhD; Abraham Assan[5], PhD; Assegid Samuel[6], MSc; Kabir Sheikh[5], PhD; Binyam Tilahun[2*], MPH, MSc, PhD

[1]Department of Health Education and Behavioural Sciences, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[2]Department of Health Informatics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[3]Department of Epidemiology and Biostatics, Institute of Public Health, College of Medicine and Health Science, University of Gondar, Gondar, Ethiopia

[4]Department of Pediatrics and Child Health, School of Medicine, College of Medicine and Health Science, University of Gondar, Gondar, Ethiopia

[5]Alliance for Health Policy and Systems Research, World Health Organization, Geneva, Switzerland

[6]Human Resources Administration Directorate, Ministry of Health, Addis Ababa, Ethiopia

[*]these authors contributed equally

Abstract

Background: A compassionate, respectful, and caring (CRC) health professional is very important for human-centered care, serving clients ethically and with respect, adhering to the professional oath, and serving as a model for young professionals. As countries try to achieve universal health coverage (UHC), quality delivery of health services is crucial. CRC health care is an initiative around the need to provide quality care services to clients and patients. However, there is an evidence gap on the status of CRC health care service delivery.

Objective: This scoping review aimed to map global evidence on the status of CRC health service delivery practice.

Methods: An exhaustive literature review and Delphi technique were used to answer the 2 research questions: “What is the current status of CRC health care practices among health workers?” and “Is it possible for health professionals, health managers, administrators, and policy makers to incorporate it into their activity while designing strategies that could improve the humanistic and holistic approach to health care provision?” The studies were searched from the year 2014 to September 2020 using electronic databases such as MEDLINE (PubMed), Cochrane Library, Web of Science, Hinari, and the World Health Organization (WHO) library. Additionally, grey literature such as Google, Google Scholar, and WorldWideScience were scrutinized. Studies that applied any study design and data collection and analysis methods related to CRC care were included. Two authors extracted the data and compared the results. Discrepancies were resolved by discussion, or the third reviewer made the decision. Findings from the existing literature were presented using thematic analysis.

Results: A total of 1193 potentially relevant studies were generated from the initial search, and 20 studies were included in the final review. From this review, we identified 5 thematic areas: the status of CRC implementation, facilitators for CRC health care service delivery, barriers to CRC health care delivery, disrespectful and abusive care encountered by patients, and perspectives on CRC. The findings of this review indicated that improving the mechanisms for monitoring health facilities, improving accountability, and becoming aware of the consequences of maltreatment within facilities are critical steps to improving health care delivery practices.

Conclusions: This scoping review identified that there is limited CRC service provision. Lack of training, patient flow volume, and bed shortages were found to be the main contributors of CRC health care delivery. Therefore, the health care system should consider the components of CRC in health care delivery during in-service training, pre-service training, monitoring and evaluation, community engagement, workload division, and performance appraisal.

(JMIR Hum Factors 2022;9(1):e30804) doi:10.2196/30804

KEYWORDS

compassionate; respectful; caring; CRC; health care delivery

Introduction

Background

Compassionate, respectful, and caring (CRC) health care delivery is an essential element for health workforces that builds a positive environment and intimacy among health care professionals, patients, and families. Worldwide, improving the quality of care using the limited skilled health workforce is a major challenge facing the health care system [1]. Universal health coverage (UHC) is realized when everyone has access to quality essential health care services with financial risk protection [2,3]. However, each year, almost half of the world's population cannot access the needed health services, and millions of people are forced into extreme poverty due to catastrophic out-of-pocket health expenditures [4].

An effective health system helps to promote UHC through the provision of equitable, quality, responsive, efficient, and resilient health services [5-7]. The health workforce is one of the 6 building blocks that make a health system function [8].

CRC health professionals are crucial to the strategy designed to improve the quality of care and which has several benefits for both the provider and end service users. Supporting a movement toward creating a CRC health workforce is in the agenda of most African countries [9-11]. In order to realize this vision, intensive effort by leaders and health managers at all levels of the health sector [12,13] as well as supporting systems and structures [14] are required.

Compassion and acting to relieve concerns, pain, distress, and suffering are fundamental to health care; these define the higher purpose of the health care system. Respect for people goes beyond accepting the notion or attitude that people have autonomous choice; rather, it is treating others in such a way that enables them to make the choice. Respecting the patient's right to self-determination, that is supporting decisions that reflect the patient's personal beliefs, values, interests, and problems, is thus central [12,15].

Even though a lot has been done globally to improve the health status of the population, it is still failing at a fundamental level [16-18]. The qualities of caring, respect, and compassion, which form the basis of care delivery and the human aspects that define it, have been replaced by a primary focus on pathways, tasks, and documentation [12,19].

Hence, this review aimed to assess CRC practices in health care delivery and highlight the reality of the situation. The goal was to identify key issues in developing guidance for health professionals, health managers, administrators, and policy makers, to inform and refine strategies that could improve the humanistic and holistic approach of health care provision.

Rationale for the Review

Ensuring CRC health services is one of the most important facilitating factors to increase access to and the continuum of care. However, challenges to implementing CRC health services persist and reflect large disparities across geographic areas and population groups [20-22]. Little is known about the status of CRC practices in the health systems and their underlying determinants.

Hence, this scoping review of the literature aimed to draw out the evidence regarding the current status of CRC health service delivery and design a CRC implementation strategy. In addition, this review of the literature aimed to identify the possible health care practices research areas that need critical insight and further investigation.

Methods

Literature Search and Search Methods

Our review aimed to identify the available evidence to provide an overview of the scoping review objectives of the current state of knowledge on CRC health service practices. This review was conducted following the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist, and it was guided by Joanna Briggs Institute (JBI) scoping review guidance [23,24]. A search was conducted for published and unpublished (grey) literature on the research area in the following electronic databases: MEDLINE (PubMed), Cochrane Library, Web of Science, Hinari, and the World Health Organization (WHO) library. Moreover, grey literature was searched on Google, Google Scholar, and WorldWideScience. All studies published from 2014 to September 2020 were considered. We used different combinations of keywords and text to build the search strategy and identify relevant articles. The searching techniques for PUBMED considered Boolean operators with the following search terms: *"((compassion*) OR compassionate OR concern) OR empathy) OR kindness) OR consideration)) AND ((Respect*) OR Respectful) OR deferent*) OR reverent*) OR polite*) OR courteous) OR considerate) OR civil)) AND ((caring) OR care OR) kind) OR thoughtful) OR gentle) OR helpful) OR considerate) OR compassionate) OR concerned) OR loving)) AND ((health professional* OR health personnel* OR health care provider* OR health care worker* OR nurses OR Midwives* OR pharmacists OR physicians OR health care worker*)) AND ((behaviour) OR performance) OR actions) OR deeds) OR activities) OR manners) OR conduct)) AND ((health care delivery OR delivery of health care OR performance OR behaviour OR health care system OR health care systems))"*.

To identify the potentially relevant literature, a hand search was also conducted of the references of the included studies and websites such as the WHO and Directory of Research on CRC. Potentially relevant grey literature was identified through targeted searches of dissertations, theses, and conference abstracts (EMBASE Conference Abstracts, Conference Proceedings Citation Index—Science and Social Science & Humanities).

Scoping Review Research Question

We used a population, concept, and context (PCC) framework developed by the JBI to determine the eligibility of our primary research question. The primary research question for this scoping review was: "What is the current status of CRC health care practices among health workers?" The other research question was "Is it possible for health professionals, health managers,

administrators, and policy makers to incorporate it into their activity while designing strategies that could improve the humanistic and holistic approach to health care provision?" This study used the PCC format (Table 1) to align the study selection with the aforementioned research question.

Table 1. Eligibility of studies according to the participant, concept, and context (PCC) framework.

Criteria	Elements
P: participants	All categories of health professionals involved in health service delivery in all levels of public, private, or other sectors of health care
C: concept	Studies that explored compassionate, respectful, and caring behaviors by health professionals in different forms were included. Compassionate care, compassion, and empathy in health care delivery, respectful health care, and caring behavior exhibited by health professionals as well as related concepts were explored in this review.
C: context	All countries in the world

Study Selection Criteria

Textbox 1 outlines the inclusion and exclusion criteria for the scoping review.

Textbox 1. Inclusion and exclusion criteria for the scoping review.

Inclusion criteria

- Focus on health care providers or health professionals
- Report on health care practices or any health care services provided to any community
- Published from 2014 to September 2020
- Qualitative and quantitative studies

Exclusion criteria

- Publication in a non-English language
- Studies for which a full-text article could not be obtained (ie, studies with no full-text were excluded after repeatedly contacting the authors)

Data Extraction and Management

Data were extracted using a standardized data extraction spreadsheet. The data extraction sheet included study characteristics such as author name, year, country, types of health services with CRC, the purpose of exercising CRC, study population, study design, and publication year. Data were extracted by 2 of the authors (AN and DAA) independently. The level of agreement between the 2 reviewers was measured using the Cohen kappa level of agreement [25]. The 2 authors resolved disagreements by discussion, consulting a third author (BFE) for any persistent disagreements.

Study Selection and Reliability

Initial searches were performed by 2 review authors with extensive experience in systematic reviews. The screening of titles, abstracts, and full texts was conducted independently by 2 review authors (AN and DAA). A disagreement regarding the decision against the inclusion of articles between the 2 reviewers was resolved by consensus or the third reviewer (BFE).

A second reviewer (DAA) was blinded to the primary reviewer's (AN) decision for checking article selection, data extraction, and risk of bias assessment stages of the reviews. Any differences of opinion were discussed; otherwise, a third reviewer (BFE) was available to arbitrate any issues that remained unresolved.

Data Analysis

The methodological framework for the scoping review was supposed to present our narrative account of findings in 2 ways [24]; first, attention was given to the basic numerical analysis of the extent and distribution of the studies included in the review. We produced distributions for the study setting geographically, by urban or rural setting, by type of publications, and by CRC health services using tables and graphs. Second, the study findings from the existing literature are presented using thematic analysis. Then, a code book (Multimedia Appendix 1) and its definitions were prepared in a separate Word document. Our narrative literature was then structured around the themes derived from the study results. The themes that emerged from the study were (1) facilitators for CRC health care delivery, (2) barriers to CRC health care delivery, (3) disrespectful and abusive care encountered by patients, and (4) perspectives on CRC health care delivery.

Results

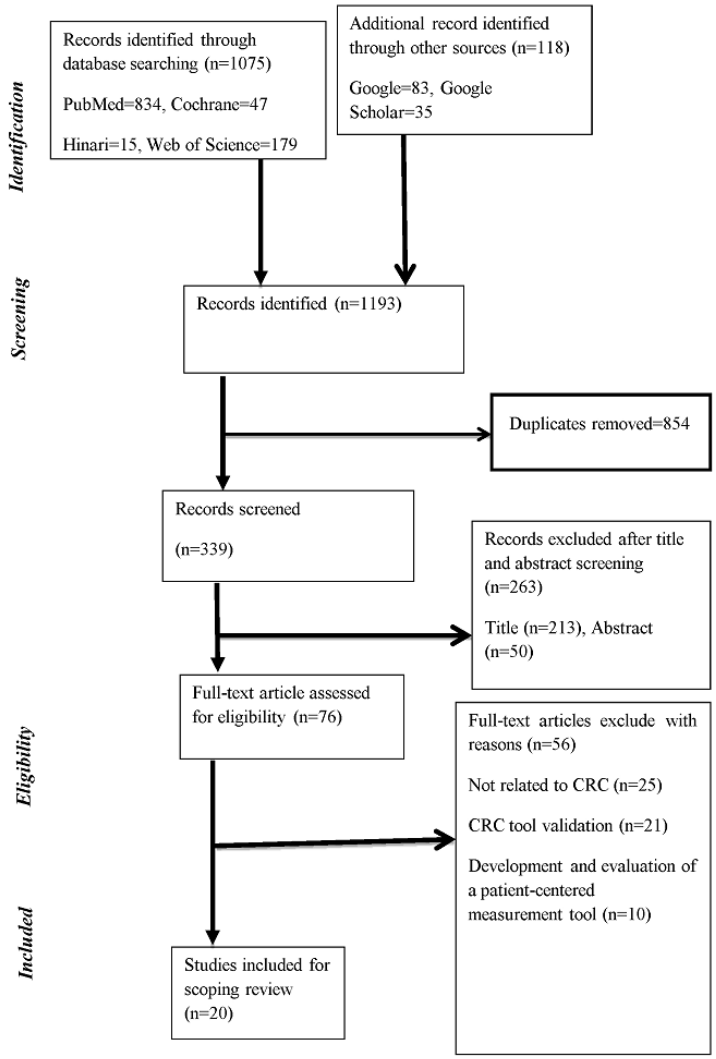
Flow of the Search and Study Characteristics

A total of 1193 potentially relevant studies (834 from PubMed, 35 from Google Scholar [advanced], 83 from Google, 47 from CINAHL, 15 from Hinari, and 179 from Web of Science) were generated from the initial search. After duplicates were excluded, 339 studies remained. Then, we excluded most of the potentially irrelevant papers based on the review of title (n=213)

and abstract (n=50). Overall, 76 studies were eligible for full-text screening. After reading the full text, 56 studies were excluded due to not being related to CRC (n=25), a focus on CRC tool validation (n=21), and a focus only on the development and evaluation of a patient-centered measurement tool (n=10); 20 articles were retained for the final review (Figure 1).

Based on the inclusion criteria, 20 studies were included in this scoping review (Multimedia Appendix 2). The idea related to compassion in health care service delivery practice was assessed and explored via a mixed (qualitatively and quantitatively) approach in 2 studies from the perspectives of health professionals and patients [26,27].

Figure 1. Flow diagram for the scoping review process adapted from the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement. CRC: compassionate, respectful, and caring.



Of the studies included in this review, 3 addressed respectful care from the perspective of both health care workers and clients, in relation to respectful maternity care, disrespect and abuse during labor and delivery services, and patient-centered communication in Burkina Faso, Nigeria, and China [28-30]. The studies showed that disrespectful and abusive health care delivery is a common challenge. Caring was also addressed in 3 studies from the perspective of health care workers and clients in Ghana, Nigeria, and Reggio Emilia, Italy. A study conducted in Nigeria showed that 93.2% of the respondents had experience with at least one form of disrespectful and abusive care [31-33].

Care and respectful care were addressed in 4 studies included in this review from the perspective of midwives, midwifery students, and clients; the studies discussed any disrespectful or abusive maternal care and respectful care during labor and childbirth from their perspective [34-37]. Denial of health care delivery, overlooking of patient-centered treatment, and low socioeconomic status were the usual problems with regard to care and respectful care. In a study conducted in Ghana, about 72% of the respondents said that maltreatment was the common problem, and 77.4% said private facilities treat their clients more respectfully than public facilities [37]. Care and compassion in health care delivery were stated in 2 studies on the principles of compassion in health care practices and from the hospitalized patient viewpoint [38,39]. Only 1 study was conducted on compassionate and respectful health care delivery and showed that the magnitude of disrespect was 43% [40].

The concepts of compassion, respect, and care were assessed in 5 studies from the perspectives of health care providers, clients, and educators of health professionals [41-45]. In this

review, a study in Ethiopia on the experience of CRC showed that 55% was good and the remaining 45% was poor. Patients' perceptions toward CRC were found to be poor in 56% of the participants [42].

Characteristics of the Included Studies

In this review, 14 studies originated from Africa, namely Ethiopia, Ghana, Nigeria, Burkina Faso, and South Africa [27-29,31,32,34-37,40-43,45], and 5 studies were from other continents, in the countries of Iran, the United Kingdom, European countries, and China [26,30,33,39,44]. A qualitative approach was used by 10 of the studies [28,29,31,33-36,41,43,44], 2 used a mixed approach [27,30], 5 were cross-sectional studies [32,37,40,42,45], 1 was a descriptive-analytical study [39], 1 was based on expert consensus [26], and 1 was an expert review [38]. Of the 20 studies, one-half (10/20, 50%) of the studies were from the perspective of health workers [26,28,29,35-37,41,43,44], 7 of the studies were from the perspectives of clients [27,30,32,33,39,40,42,45], and the other 3 studies were from both health workers' and clients' perspectives [31,34] (Table 2). The studies included in this review were published between 2014 [28] and 2020 [45].

Of the 20 studies, 8 were about the services delivered for maternal health components [28,32-36,40,45], 11 studies were about the services given for all clients [26,29-31,37-39,41-44], and 1 study was with oncology patients [27].

Table 2. Summary of studies included in the scoping review of compassionate, respectful, and caring (CRC) health care delivery (2014 to 2020).

Variable	Results, n (%)
Perspectives	
Health professional	10 (50)
Client	7 (35)
Both	3 (15)
Publication	
Original article	18 (90)
Expert review	2 (10)

Themes in the Studies

Based on the review of the articles, the primary focus of the included studies fitted broadly into 5 key themes, namely (1) the status of CRC implementation, (2) facilitators for CRC health care delivery, (3) barriers to CRC health care delivery, (4) disrespectful and abusive care encountered by patients, and (5) perspectives on CRC.

Status of CRC Implementation

Compassionate care is about empathetic conversations with the client through an evidence- based framework for the safest and most trustful interaction. This evidence-based framework takes into account the transcultural settings and needs. Health care is provided for any individual who is in need of the services in a good, compassionate manner for all ages and all illnesses, taking into account the transcultural settings and the needs, wishes, and expectations of most of the users [26]. Treating the patient as a person, not just a disease, and listening attentively to patients are the elements of compassionate care with the highest rating [27]. In this review, the status of compassionate and respectful maternity care and associated factors in health facility–based childbirth was identified, and the prevalence of experienced compassionate and respectful maternity care was 57% [40].

Patient-centered care was examined with the elements of a patient-centered care framework: respecting patients' values, access to care, information access, patient empowerment, family and friend involvement, emotional support, and continuity of care [46]. This review showed that patients experienced at least one form of disrespect and abuse during health care service utilization [32,45].

Facilitators for CRC Health Care Delivery

There were a number of facilitators that enabled the health care delivery system by integrating the elements of CRC into every component of health care. Shared decision-making, fundamental principles of compassionate care, knowledge and technical skill building, in-service education, monitoring and accountability, sociodemographic characteristics of individuals, and ensuring comprehensiveness of care are some of the facilitators coded under these themes.

Shared Decision-making

Health care is provided for any individual who is in need of the services and in a compassionate manner for all ages and all illnesses, taking into account the transcultural settings and the needs, wishes, and expectations of most of the users. Considering shared decision-making as a cornerstone of evidence-based health care practice can be examined as moving beyond simply having empathetic conversations to developing a compassionate, evidence-based framework for safe and trusting interaction with the client [26]. "Treat you as a person, not just a disease, and listen attentively to you" were elements of compassionate care with the highest rating [27].

Fundamental Principles of Compassionate Care

This review found that the fundamental principles of compassionate care, such as trust, dignity, and respect, as well as effective communication skills and collaboration with patients and their families, are core requirements for the essential skills cluster "Care, Compassion, and Communication" and are required for the delivery of high-quality care [38]. All health professionals working across all health care settings should be closely aware of the concept of compassionate care.

Knowledge and Technical Skill Building, as Well as In-Service Education

Newly qualified health professionals must acquire sufficient knowledge and technical skills to care for patients and develop and demonstrate the attitudes and interpersonal attributes that characterize compassionate care [30].

Nurses need to see the patient's care needs and expectations from the patient's point of view and pay more attention to the aspects that are more important for the patients. Paying more attention to compassionate nursing care in nursing textbooks is recommended, and nurses should receive in-service education in this regard [39]. The status of compassionate and respectful maternity care and associated factors in health facility–based childbirth was identified, and the results indicated that we should do more on the issue [40].

Clients expressed moderate enthusiasm for patient-centered communication and strong preferences concerning physician respect for the patient perspective but less concern for power sharing. This means that patients were more concerned about doctors exhibiting a caring perspective than power sharing. To have respectful health care delivery services, health care providers need training on how to incorporate elements of respectful maternity care into practice, including skills for rapport building and counseling [29].

Monitoring, Accountability, and Sociodemographic Characteristics.

Frequent monitoring of care provision in health care facilities is needed to eliminate the incidence of disrespectful and abusive care. Midwives have described disrespectful and abusive care as the provision of inadequate care, overlooking of patient-centered care, and verbal, physical, and psychological abuse. Disrespectful and abusive care is facilitated by socioeconomic inequalities, provider perception, victim blaming, and other health system–related factors [35].

The provision of high-quality, patient-centered, respectful care for all patients, including laboring women, by health care professionals who were committed to providing a safe birthing experience for the patients and believed that yelling, shouting, and even hitting women in order to ensure a positive outcome was justified, understood, and perhaps even appreciated by women [36].

Maltreatment is a problem in the utilization of health care services. Private health facilities treated their patients more respectfully than public facilities. The majority of midwifery students throughout Ghana witnessed disrespectful care during their training. Improving monitoring, accountability, and consequences for maltreatment within facilities to improve the care that pregnant and laboring women receive is very important to having respectful care practices [37].

Ensuring Comprehensiveness of Care

Patient-centered care was perceived as providing quality care, making partnerships, provision of information, patient involvement, and understanding patient preference. Patient empowerment and family and friend involvement in patient care were found to be far from the existing practice and were less common in the presence of low patient health literacy levels. Patient-centered care was examined with the elements of a patient-centered care framework: respecting patients' values, access to care, information access, patient empowerment, family and friend involvement, emotional support, and continuity of care [31].

The development of a feasible multicomponent palliative care intervention by involving clients in the decision-making process and conducting appropriate educational processes and training for health workers was significant for meeting the needs of the clients [33].

Barriers to CRC Health Care Delivery

In this review, we identified a number of barriers that hinder the integration of CRC health care delivery. Some of the codes comprising the subthemes include a lack of human resource development, lack of infrastructure at the health facility, poor behavior by health workers, poor client experience at the health facility, and client sociodemographic characteristics [28-31,34,35,40,41].

From the perspective of health care professionals, the most commonly cited barriers to CRC health care delivery were busy staff, absence of close follow-ups or monitoring by the leaders; knowledge and attitude gaps; absence of an information desk, complaint handling mechanism, a mechanism to obtain client feedback, discrimination-free care, friendly care, regular health education in the health facility, and capacity building; high patient flow; bed shortages; being treated by different physicians; high staff turnover; and large numbers of patients in referral centers [27,28,30,41]. From the clients' perspectives, the most commonly cited barriers to CRC care were abuse of clients during facility health service delivery, violation of clients' rights, occupational status of the client, previous experience with health facility utilization, residency, time, complication during service delivery, family monthly income, and intention to use a health facility [40,45].

Disrespect and Abusive Care Encountered by Patients

In this review, there were many forms of disrespectful and abusive care encountered by patients or clients that need to be addressed properly. The most common forms of disrespect and abuse experienced by clients were nondignified care, lack of privacy, physical abuse, neglectful care, nonconfidential care, and nonconsented care. Clients have suggested that more attention should be paid to patients' care needs, expectations from the patient's point of view, and the aspects that are more important for the patients [32,33,39,40,45].

The nature of disrespect and abuse in midwifery care during labor and delivery involves the denial of the preferred birth position, denial of accompaniment, denial of care, poor clinical practice, neglect, and verbal abuse from the patients' perspectives, as well as denial of service; verbal abuse; physical abuse like hitting, pinching, and slapping; and violation of privacy from the midwives' perspectives and women's experiences with care from midwives during labor and delivery, including any disrespect or abuse [34].

Perspectives on CRC

Perspective is an art technique that changes the space or depth of CRC on health care delivery. The way of looking at CRC is not similar for everyone (ie, the health care provider and client may look at CRC in different ways). In this review, health care providers and clients were looking at CRC health care delivery in different ways, which is very important to take into account to develop an appropriate intervention.

Health Care Provider Perspectives on CRC

Great emphasis on and recommendations for education (CRC attitude and value more than cognitions) are needed. Designing an educational curriculum with respect to CRC for all health professionals in a higher educational program from which health professionals will graduate is the best strategy to produce health care workers who could practice his or her profession in the appropriate and ethical way [37,43]. “Caring” was taken to mean being able to converse well, up to-date, and proficient in the field of work as well as being considerate and respectful to others. Professional midwives indicated that they have seen colleagues demonstrate uncaring behavior, educators emphasized respect as caring, and student midwives, professional midwives, and educators described caring as being a competent nurse with compassion and respect for others [43].

Positive achievements in CRC health care delivery were feedback to colleagues who did not follow the training recommendations, commitment of trained health workers, improved cleanliness in the delivery room, ample explanation given to the client before providing a diagnosis, and communication with the parturient before and during the intervention [28].

Disrespect and abuse in health care and the impact on the health and well-being of the patient were perpetrated or witnessed as a violation of human rights while highlighting the patient’s expectations of care as the basis for the subjectivity of experiences [29].

A study in the Tigray region showed that the experience of patients with CRC health care practice was reported as good by 55% of respondents. In contrast, patients’ perceptions toward CRC was found to be poor, as reported by 56% of the participants [42].

Compassionate care from the perspective of staff working in health settings wishing to provide compassionate care, on its own, was insufficient to ensure this transpired; health care providers needed to work in a setting that supports them doing this, which underpins our core concept of the compassionate care flow. As “professional” compassion was associated with the intention to improve patient health and participants’ roles within health care, a compassionate care flow could be enhanced by defenders (eg, supportive colleagues, seeing the patient as a person, drawing on their faith) or depleted by drainers (ie, competing demands on time and resources), through their impact on professional compassion [44].

Client Perspectives on CRC

Clients who visit a health facility for different health care services stated that compassion and respectful health care services provided by health personnel encouraged them to visit again for any other medical check-up. Many of the clients wanted to be treated by one physician based on their choice, but most of the clients reported that they were treated by different physicians without being asked their choice [27].

Clients expressed strong preferences concerning physician respect for patient perspectives rather than for power sharing. Younger and highly educated patients were more likely to prefer patient-centered communication, and highly educated patients paid more attention to power sharing [30].

Discussion

Principal Findings

This systematic scoping review explored the available literature on CRC health care delivery, based on the perspectives of health professionals and patients or clients in the global context. In this review, as a country, Ethiopia was the most represented, which might be linked with the fact that Ethiopia initiated the design of a 5-year Compassionate, Respectful and Caring Health Services Implementation Strategy as one of the national top priorities set under the health sector transformation plan [12]. In this review, 90% (18/20) of all articles were published in the last 5 years, suggesting that patients or clients and health professionals increasingly view patient-centered communication, compassion, respect, and caring as essential for good health service utilization [26,29-31,33,39,43-45].

In addition, the current review delivers some insight into the status of compassion, suggesting that compassionate care is about empathetic conversations with the client through an evidence-based framework for the safest and trustful interaction and treating patients as a person, not just a disease, and listening attentively were elements of compassionate care with the highest rating [26,27].

CRC health care delivery behaviors have direct impacts on health-seeking behaviors and overall health outcomes [47]. Compassion does not depend on pre-existing relationships; rather, it is delivered through a long-term relationship with individuals. Compassion consists of specific skills and actions aimed at the enhancement of multifactorial suffering, namely, acknowledging, responding to, understanding, and actively addressing the suffering of another [47].

This review on the concepts of respectful health care delivery showed that patients were more concerned about doctors exhibiting caring perspectives than power-sharing. There was more emphasis on patient-centered communication and strong preferences concerning physician respect for patient perspectives and less concern for power-sharing [30]. High staff turnover and large numbers of patients in referral centers were the main challenges observed for respectful health care delivery [29].

Disrespectful and abusive practices were witnessed as a violation of human rights while highlighting women’s expectations of care as the basis for the subjectivity of experiences [29]. Respectful health care is the factor most neglected in health care provision [48] but that increases client satisfaction and affects the health-seeking behaviors of the community. A

practice of respectful health care is a set of safe health facilities for health care services where the clients or service users are valued, recognized, treated fairly, have clear expectations, and appropriately access services, as needed. When there is a practice of respectful health care delivery at the health institution, the likelihood of the community using the facility could be increased in a remarkable way [49].

A patient-centered framework is an essential element for the patient, one that is guided by respect for the patient's values, access to care, access to information, patient empowerment, involvement of family and friends, emotional support, and continuity of care [12]. Engagement of clients influences both the overall health care services and improves health care provision because both clients and health care providers feel respected, listened to, and empowered [50,51].

CRC health care delivery is essential for the successful utilization of health services. The findings from this review suggest that CRC health care delivery might have a positive effect on specific health care service utilization, including increased service satisfaction and sustainability of service utilization. The barriers toward implementation of CRC health care practice in this review from the perspective of health care professionals are busy staff; absence of close follow-ups or monitoring by the leaders; knowledge and attitude gaps; and absence of an information desk, a complaint handling mechanism, mechanism for obtaining client feedback, and regular health education in a health facility [41]. This is in agreement with a review conducted in a clinical health care setting [52].

In this review, in a study on caring from the perspectives of undergraduate student midwives, professional midwives, and teachers of midwifery, participants described caring as being a competent nurse with compassion and respect for others [43]. This finding could help to design an integrated curriculum for health professional teaching and to have a strong CRC norm in health care delivery. Similarly, a cross-sectional study conducted in Ethiopia on the status of CRC with perspective clients showed that the experience and perception of patients toward CRC health care practice were good, which means that clients had experience with CRC health care service from a health care provider [42]. On the other hand, there are some barriers for the implementation of CRC in health care practices such as busy staff due to overloaded work; absence of close follow-ups or monitoring by the leaders; knowledge and attitude gaps; and absence of an information desk, a complaint handling mechanism, mechanism for obtaining client feedback, regular health education in a health facility, and capacity building. Therefore, we recommend that enhancing CRC health care delivery in the health care system requires empirical teaching methods as a baseline that engage the learner professionally, ethically, and personally, because compassion, respect, and caring are rooted in the nature of the students and the actualization of these qualities within health care service delivery practices, in addition to considering periodical refresher and follow-up training for those who are employed in the health care delivery system.

Strengths and Limitations

Scoping reviews are broad in nature and provide an overview of existing literature regardless of quality, providing a broader and more contextual overview than systematic reviews. Use of the PRISMA-ScR was a strength of this scoping review. A formal assessment of methodological quality is not undertaken when conducting a scoping review and synthesis of the incorporated studies. Including papers published only in English was another limitation of this review.

Conclusion

This scoping review showed the status of disrespectful and abusive care encountered by the patients, facilitators of and barriers to CRC health care delivery behaviors, and different perspectives. The status of CRC health care delivery remains challenging and needs strong involvement from different organizations from all disciplines. Pre-service education with full CRC competencies (ie, higher education institutes) should include CRC as one component of the curriculum for health professionals, as well as frequent in-service training and inclusion of CRC elements in human resource selection, performance management, and incentive systems, including career advancement, deployment opportunities, and labor division at health facilities.

Acknowledgments

The authors acknowledge the University of Gondar, the World Health Organization (WHO), and authors of the included studies. This work was supported by the Alliance for Health Policy and Systems Research. The Alliance is able to conduct its work thanks to the commitment and support from a variety of funders. These include long-term core contributors from national governments and international institutions, as well as designated funding for specific projects within the Alliance's current priorities.

Authors' Contributions

BT, AN, BFE, and DAA were involved in the conception and design of this review. BT guided the manuscript development. AN, BT, BFE, and DAA participated in the writing of the final manuscript. BT, ZAM, AT, MF, AA, AS, and KS reviewed the final manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Codebook for thematization of the study.
[DOCX File , 22 KB - humanfactors_v9i1e30804_app1.docx]

Multimedia Appendix 2
Study characteristics.
[DOCX File , 29 KB - humanfactors_v9i1e30804_app2.docx]

References

1. Clift C. Is Universal Health Coverage Good for Health? Universal Health Coverage and the post-2015 development agenda. Research Gate. 2013. URL: https://www.researchgate.net/publication/291102850 [accessed 2022-01-22]
2. What is universal health coverage? World Health Organization. 2020 Jan 20. URL: https://www.who.int/news-room/questions-and-answers/item/what-is-universal-health-coverage [accessed 2022-01-22]
3. Monitoring progress towards universal health coverage at country and global levels. World Health Organization. 2014 May. URL: http://apps.who.int/iris/bitstream/10665/112824/1/WHO_HIS_HIA_14.1_eng.pdf [accessed 2022-01-22]
4. Hogan DS, Stevens G, Hosseinpoor A, Boerma T. Monitoring universal health coverage within the Sustainable Development Goals: development and baseline data for an index of essential health services. The Lancet Global Health 2018 Feb 13;6(2):e152-e168 [FREE Full text] [doi: 10.1016/s2214-109x(17)30472-2]
5. Everybody's business: Strengthening health systems to improve health outcomes. World Health Organization. 2007. URL: https://www.who.int/healthsystems/strategy/everybodys_business.pdf [accessed 2022-01-22]
6. The World health report: 2000: Health systems: improving performance. World Health Organization. URL: https://apps.who.int/iris/handle/10665/42281 [accessed 2022-01-22]
7. Measuring Up: Improving Health System Performance in OECD Countries. OECD. 2002. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/measuring-up_9789264195950-en [accessed 2022-01-22]
8. Health systems. World Health Organization. URL: https://www.euro.who.int/en/health-topics/Health-systems [accessed 2022-01-22]
9. Chochinov HM. Dignity in care: time to take action. J Pain Symptom Manage 2013 Nov;46(5):756-759 [FREE Full text] [doi: 10.1016/j.jpainsymman.2013.08.004] [Medline: 24099895]
10. Lown B, Rosen J, Marttila J. An agenda for improving compassionate care: a survey shows about half of patients say such care is missing. Health Aff (Millwood) 2011 Sep;30(9):1772-1778 [FREE Full text] [doi: 10.1377/hlthaff.2011.0539] [Medline: 21900669]
11. Chochinov HMC. Dignity and the essence of medicine: the A, B, C, and D of dignity conserving care. BMJ 2007 Jul 28;335(7612):184-187 [FREE Full text] [doi: 10.1136/bmj.39244.650926.47] [Medline: 17656543]
12. Health Sector Transformation Plan II. Ministry of Health - Ethiopia. 2021 Feb. URL: https://fp2030.org/sites/default/files/HSTP-II.pdf [accessed 2022-01-22]
13. Rodriguez A, Lown B. Measuring compassionate healthcare with the 12-item Schwartz Center Compassionate Care Scale. PLoS ONE 2019 Sep 5;14(9):e0220911 [FREE Full text] [doi: 10.1371/journal.pone.0220911]
14. West M, Armit K, Loewenthal L, Eckert R, West T, Lee A. Leadership and leadership development in health care: the evidence base. The King's Fund. 2015. URL: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/leadership-leadership-development-health-care-feb-2015.pdf [accessed 2022-01-22]
15. Smeltzer SC, Bare BG, Hinkle JL, Cheever KH. Brunner & Suddarth's Textbook of Medical-Surgical Nursing. Philadelphia, PA: Lippincott Williams & Wilkins; 2010.
16. Peabody JW, Taguiwalo MM, Robalino DA, Frenk J, Jamison DT, Breman JG, et al, editors. Improving the Quality of Care in Developing Countries. In: Disease Control Priorities in Developing Countries. Washington, DC: The International Bank for Reconstruction and Development / The World Bank; 2006.
17. Delivering quality health services: a global imperative for universal health coverage. World Health Organization. 2018. URL: https://apps.who.int/iris/handle/10665/272465 [accessed 2022-01-22]
18. Chisholm D, Evans DB. Improving health system efficiency as a means of moving towards universal coverage. World Health Organization. 2010. URL: https://www.who.int/healthsystems/topics/financing/healthreport/28UCefficiency.pdf [accessed 2022-01-22]
19. Smiley S, King J, Hancock WW. Fibrinogen stimulates macrophage chemokine secretion through toll-like receptor 4. J Immunol 2001 Sep 01;167(5):2887-2894 [FREE Full text] [doi: 10.4049/jimmunol.167.5.2887] [Medline: 11509636]
20. Henry K, Niu X, Boscoe FP. Geographic disparities in colorectal cancer survival. Int J Health Geogr 2009 Jul 23;8(1):8-48 [FREE Full text] [doi: 10.1186/1476-072x-8-48]
21. Govindarajan V, Ramamurti R. Delivering world-class health care, affordably. Harvard Business Review. 2013. URL: https://hbr.org/2013/11/delivering-world-class-health-care-affordably [accessed 2022-01-22]

22. Farr M, Barker R. Can staff be supported to deliver compassionate care through implementing Schwartz rounds in community and mental health services? Qual Health Res 2017 Sep;27(11):1652-1663 [FREE Full text] [doi: 10.1177/1049732317702101] [Medline: 28799475]
23. Peters M, Godfrey C, Khalil H, McIrnerney P, Soares C, Parker D. Guidance for the Conduct of JBI Scoping Reviews. In: Aromataris E, Munn Z, editors. Joanna Briggs Institute Reviewer's Manual. Adelaide, South Australia: The Joanna Briggs Institute; 2017.
24. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
25. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012:276-282 [FREE Full text] [doi: 10.11613/bm.2012.031]
26. Mechili E, Angelaki A, Petelos E, Sifaki-Pistolla D, Chatzea V, Dowrick C, et al. Compassionate care provision: an immense need during the refugee crisis: lessons learned from a European capacity-building project. J of Compassionate Health Care 2018 Jan 12;5(1):1-8 [FREE Full text] [doi: 10.1186/s40639-018-0045-7]
27. Zeray M. Compassionate care among oncology patients at Tikur Anbssa Specialized Hospital: tool validation and measurement of practice in Addis Ababa, Ethiopia. AAU Institutional Repository. 2018. URL: http://etd.aau.edu.et/handle/123456789/14549 [accessed 2022-01-22]
28. Ouédraogo A, Kiemtoré S, Zamané H, Bonané BT, Akotionga M, Lankoande J. Respectful maternity care in three health facilities in Burkina Faso: the experience of the Society of Gynaecologists and Obstetricians of Burkina Faso. Int J Gynaecol Obstet 2014 Oct 18;127 Suppl 1:S40-S42 [FREE Full text] [doi: 10.1016/j.ijgo.2014.07.009] [Medline: 25087176]
29. Orpin J, Puthussery S, Burden B. Healthcare providers' perspectives of disrespect and abuse in maternity care facilities in Nigeria: a qualitative study. Int J Public Health 2019 Dec 31;64(9):1291-1299 [FREE Full text] [doi: 10.1007/s00038-019-01306-0] [Medline: 31673736]
30. Ting X, Yong B, Yin L, Mi T. Patient perception and the barriers to practicing patient-centered communication: A survey and in-depth interview of Chinese patients and physicians. Patient Educ Couns 2016 Mar;99(3):364-369 [FREE Full text] [doi: 10.1016/j.pec.2015.07.019] [Medline: 26776708]
31. Nkrumah J, Abekah-Nkrumah G. Facilitators and barriers of patient-centered care at the organizational-level: a study of three district hospitals in the central region of Ghana. BMC Health Serv Res 2019 Nov 27;19(1):900 [FREE Full text] [doi: 10.1186/s12913-019-4748-z] [Medline: 31775809]
32. Ogunlaja A, Fehintola O, Ogunlaja I, Popoola G, Idowu A, Awotunde O, et al. "Respectful Maternity Care" or "Disrespect and Abuse during Maternity Care"; Experience of pregnant women in Ogbomoso, South West Nigeria. Rwanda Med J 2017;74(3):6-9 [FREE Full text] [doi: 10.24125/sanamed.v11i3.152]
33. Tanzi S, Alquati S, Martucci G, De Panfilis L. Learning a palliative care approach during the COVID-19 pandemic: A case study in an Infectious Diseases Unit. Palliat Med 2020 Oct;34(9):1220-1227 [FREE Full text] [doi: 10.1177/0269216320947289] [Medline: 32736486]
34. Burrowes S, Holcombe S, Jara D, Carter D, Smith K. Midwives' and patients' perspectives on disrespect and abuse during labor and delivery care in Ethiopia: a qualitative study. BMC Pregnancy Childbirth 2017 Aug 22;17(1):263 [FREE Full text] [doi: 10.1186/s12884-017-1442-1] [Medline: 28830383]
35. D-zomeku V, Mensah B, Nakua E, Agbadi P, Lori J, Donkor P. Exploring midwives' understanding of respectful maternal care in Kumasi, Ghana: Qualitative inquiry. PLoS ONE 2020 Jul 9;15(7):e0220538 [FREE Full text] [doi: 10.1371/journal.pone.0220538]
36. Rominski S, Lori J, Nakua E, Dzomeku V, Moyer CA. When the baby remains there for a long time, it is going to die so you have to hit her small for the baby to come out": justification of disrespectful and abusive care during childbirth among midwifery students in Ghana. Health Policy Plan 2017 Mar 01;32(2):215-224 [FREE Full text] [doi: 10.1093/heapol/czw114] [Medline: 28207054]
37. Moyer C, Rominski S, Nakua E, Dzomeku V, Agyei-Baffour P, Lori JR. Exposure to disrespectful patient care during training: Data from midwifery students at 15 midwifery schools in Ghana. Midwifery 2016 Oct;41:39-44 [FREE Full text] [doi: 10.1016/j.midw.2016.07.009] [Medline: 27522042]
38. Bloomfield J, Pegram A. Organisational aspects of care. Nurs Stand 2015 Mar 04;29(27):35-40 [FREE Full text] [doi: 10.7748/ns.29.27.35.e9581] [Medline: 25736672]
39. Dalvandi A, Vaisi-Raygani A, Nourozi K, Ebadi A, Rahgozar M. The importance and extent of providing compassionate nursing care from the viewpoint of patients hospitalized in educational hospitals in Kermanshah - Iran 2017. Open Access Maced J Med Sci 2019 Mar 30;7(6):1047-1052 [FREE Full text] [doi: 10.3889/oamjms.2019.204] [Medline: 30976359]
40. Wassihun B, Zeleke S. Compassionate and respectful maternity care during facility based child birth and women's intent to use maternity service in Bahir Dar, Ethiopia. BMC Pregnancy Childbirth 2018 Jul 09;18(1):294 [FREE Full text] [doi: 10.1186/s12884-018-1909-8] [Medline: 29986659]
41. Berhe H, Berhe H. Barriers towards implementation of caring, respectful and compassionate healthcare practice in Tigray regional state, Ethiopia: a qualitative study. Research & Reviews: Journal of Medical Science and Technology 2019;7(3):33-38 [FREE Full text]

42. Berhe H, Berhe H, Bayray A, Godifay H, Gigar G, Beedemariam G. Status of caring, respectful and compassionate health care practice in Tigrai regional state: patients' perspective. International Journal of Caring Sciences 2017;10(3):1118-1128 [FREE Full text]
43. Masala-Chokwe MET, Ramukumba TS. Exploring the meaning of caring amongst student midwives, professional midwives and educators in Tshwane, South Africa. African Journal of Primary Health Care & Family Medicine 2015;7(1):a894 [FREE Full text] [doi: 10.4102/phcfm.v7i1.894]
44. Tierney S, Seers K, Tutton E, Reeve J. Enabling the flow of compassionate care: a grounded theory study. BMC Health Serv Res 2017 Mar 03;17(1):174 [FREE Full text] [doi: 10.1186/s12913-017-2120-8] [Medline: 28253874]
45. Bekele W, Bayou N, Garedew MG. Magnitude of disrespectful and abusive care among women during facility-based childbirth in Shambu town, Horro Guduru Wollega zone, Ethiopia. Midwifery 2020 Apr;83:102629 [FREE Full text] [doi: 10.1016/j.midw.2020.102629] [Medline: 31981935]
46. Bogale T, Beharu M, Tesfaye T, Belay Y. Scope of patient centered care practice in public hospitals of Benishangul Gumuze regional state, south West Ethiopia. Quality in Primary Care 2018;26(1):31-37 [FREE Full text] [doi: 10.11648/j.rs.20210601.12]
47. Vivino B, Thompson B, Hill C, Ladany N. Compassion in psychotherapy: the perspective of therapists nominated as compassionate. Psychother Res 2009 Mar;19(2):157-171. [doi: 10.1080/10503300802430681] [Medline: 19034730]
48. Reader TW, Gillespie A. Patient neglect in healthcare institutions: a systematic review and conceptual model. BMC Health Serv Res 2013 Apr 30;13(1):156 [FREE Full text] [doi: 10.1186/1472-6963-13-156] [Medline: 23631468]
49. Restubog SLD, Hornsey MJ, Bordia P, Esposo SR. Effects of psychological contract breach on organizational citizenship behaviour: Insights from the group value model. Journal of Management Studies 2008;45(8):1377-1400 [FREE Full text] [doi: 10.1111/j.1467-6486.2008.00792.x]
50. The world health report 2008: primary health care now more than ever. World Health Organization. 2008. URL: https://apps.who.int/iris/handle/10665/43949 [accessed 2022-01-22]
51. Boivin AC, Currie K, Fervers B, Gracia J, James M, Marshall C, G-I-N PUBLIC. Patient and public involvement in clinical guidelines: international experiences and future perspectives. Qual Saf Health Care 2010 Oct;19(5):e22. [doi: 10.1136/qshc.2009.034835] [Medline: 20427302]
52. Sinclair SN, Norris JM, McConnell SJ, Chochinov HM, Hack TF, Hagen NA, et al. Compassion: a scoping review of the healthcare literature. BMC Palliat Care 2016 Jan 19;15:6 [FREE Full text] [doi: 10.1186/s12904-016-0080-0] [Medline: 26786417]

Abbreviations

CRC: compassionate, respectful, and caring
JBI: Joanna Briggs Institute
PCC: participant, concept, and context
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
UHC: universal health coverage
WHO: World Health Organization

Edited by A Kushniruk; submitted 29.05.21; peer-reviewed by S Hertling, O Aldiabat, S Pesälä; comments to author 23.07.21; revised version received 03.09.21; accepted 09.10.21; published 07.02.22.

Please cite as:
Nigusie A, Endehabtu BF, Angaw DA, Teklu A, Mekonnen ZA, Feletto M, Assan A, Samuel A, Sheikh K, Tilahun B
Status of Compassionate, Respectful, and Caring Health Service Delivery: Scoping Review
JMIR Hum Factors 2022;9(1):e30804
URL: https://humanfactors.jmir.org/2022/1/e30804
doi:10.2196/30804
PMID:35129450

Review

Factors Associated With Willingness to Share Health Information: Rapid Review

Iffat Naeem[1], MSc; Hude Quan[2], PhD; Shaminder Singh[3], PhD, RN; Nashit Chowdhury[2], MBBS; Mohammad Chowdhury[4], PhD; Vineet Saini[1,2,5], PhD, DVM; Turin TC[2,6], MBBS, PhD

[1]O'Brien Institute of Public Health, University of Calgary, Calgary, AB, Canada

[2]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3]School of Nursing and Midwifery, Faculty of Health, Community and Education, Mount Royal University, Calgary, AB, Canada

[4]Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[5]Research and Innovation – Provincial Population and Public Health, Alberta Health Services, Calgary, AB, Canada

[6]Department of Family Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Iffat Naeem, MSc
O'Brien Institute of Public Health
University of Calgary
GD01, TRW Building
3280 Hospital Drive NW
Calgary, AB, T2N4N1
Canada
Phone: 1 5877787866
Email: Iffat.naeem1@ucalgary.ca

Abstract

Background: To expand research and strategies to prevent disease, comprehensive and real-time data are essential. Health data are increasingly available from platforms such as pharmaceuticals, genomics, health care imaging, medical procedures, wearable devices, and internet activity. Further, health data are integrated with an individual's sociodemographic information, medical conditions, genetics, treatments, and health care. Ultimately, health information generation and flow are controlled by the patient or participant; however, there is a lack of understanding about the factors that influence willingness to share health information. A synthesis of the current literature on the multifactorial nature of health information sharing preferences is required to understand health information exchange.

Objective: The objectives of this review are to identify peer-reviewed literature that reported factors associated with health information sharing and to organize factors into cohesive themes and present a narrative synthesis of factors related to willingness to share health information.

Methods: This review uses a rapid review methodology to gather literature regarding willingness to share health information within the context of eHealth, which includes electronic health records, personal health records, mobile health information, general health information, or information on social determinants of health. MEDLINE and Google Scholar were searched using keywords such as *electronic health records* AND *data sharing* OR *sharing preference* OR *willingness to share*. The search was limited to any population that excluded health care workers or practitioners, and the participants aged ≥18 years within the US or Canadian context. The data abstraction process using thematic analysis where any factors associated with sharing health information were highlighted and coded inductively within each article. On the basis of shared meaning, the coded factors were collated into major themes.

Results: A total of 26 research articles met our inclusion criteria and were included in the qualitative analysis. The inductive thematic coding process revealed multiple major themes related to sharing health information.

Conclusions: This review emphasized the importance of data generators' viewpoints and the complex systems of factors that shape their decision to share health information. The themes explored in this study emphasize the importance of trust at multiple levels to develop effective information exchange partnerships. In the case of improving precision health care, addressing the factors presented here that influence willingness to share information can improve sharing capacity for individuals and allow researchers to reorient their methods to address hesitation in sharing health information.

(JMIR Hum Factors 2022;9(1):e20702) doi:10.2196/20702

KEYWORDS

health information; information sharing; health data; EMR; PHR; mobile phone

Introduction

Background

In the age of precision medicine and precision public health, good-quality data are an imperative first step to inform clinical guidelines, best practices, and policies. Precision medicine focuses on individualized patient care, taking into account the variability in genes, environment, and lifestyle [1]. Precision public health emphasizes targeted intervention programs for disease prevention and health promotion to reduce health disparities in populations [2]. This is done by applying emerging methodologies in epidemiology, biostatistics, and computing systems, including machine learning and artificial intelligence. These concepts often intersect, where clinical guidelines developed from population-level studies are adjusted to an individual patient based on their unique characteristics, leading to optimal care [3].

To expand research, and medical and disease prevention agendas through the use of precision medicine and public health frameworks, comprehensive, real-time, and integrated data need to be available. A growing number of diverse data sources provide rich and complex information, including pharmaceuticals, genomics, health care imaging, medical procedures, wearable devices, and internet activity [4]. The potential to harness these data to inform health care systems and health delivery is vast. This can include research on large, shared medical data sets or population-level sources to formulate disease risk models for use at the point-of-care level and to inform precision policy [5]. Further, eHealth data are increasingly available and provide integrated information about an individual's sociodemographic information, medical conditions, genetics, treatments, and health care use [6]. Collection, aggregation, analysis, and dissemination of electronically stored individualized health information allow for an opportunity to have a more integrated and coordinated health care system [7].

Multiple forms of health information exchange can occur: (1) health information sharing between health agencies, (2) individuals sharing health data with medical care providers, and (3) individuals sharing health data in health research studies, health social networks, biobanks, and nationwide health information exchanges [8]. Ultimately, the patient or participant is at the center of this information exchange network; therefore, understanding the willingness, interest, and motivation to provide health information is an important aspect that must be explored [9]. Willingness to share information pertains to the *intention* to perform the sharing behavior and can be defined as the extent to which a person is ready to share their intellectual capital with other individuals. Willingness may be viewed as a mediator between the factors that influence sharing health information, which make up a sharer's cognitive thought process, and the act of sharing. The concept of the intention (or willingness) to share precedes the sharing behavior following the theory of planned behavior, as posited by Madden et al [10]. This theoretical framework outlines that the attitudes, subjective norms, and the perceptions of control a person has, influences the intention to perform a behavior. In the case of willingness to share health information, willingness to share may be dependent on the perception of that individual regarding how favorable or unfavorable the result of sharing would be [11]. In this case, willingness to share health information may be a careful weighing of factors that may operate as positive or negative to influence a person to contribute their information.

National surveys and eHealth information platforms can provide excellent opportunities to collect health information for research and surveillance but can only be done if they are shared by the individuals being questioned. Previous studies have reported a high proportion of participants willing to share their health information for multiple purposes (care improvement, research, or surveillance) [8,12,13]. Privacy concerns and the type of information shared are considered important factors in studies understanding sharing preferences among patients sharing information toward electronic health records (EHRs) or personal health records. However, sharing of health information is nuanced by the influence of multiple factors, which can include information security, uncertainty about the end use of information, altruism, personality traits, illness histories, and other attributes related to the context around information sharing [14].

To our knowledge, there is no synthesis of studies that summarize the factors that individuals consider when sharing their health information. A synthesis of the current literature might help us to be able to link the various correlates of health information sharing preferences to ultimately increase the data sharing potential in certain populations. This rapid review offers an alternative form of knowledge synthesis compared with systematic review, where the process of review conduction is simplified and result synthesis can be done in a timely fashion. The results of the rapid review are usually descriptive and provide readily available knowledge about a topic in order to inform further investigation and decision-making [15]. For the purposes of this study, conducting a rapid review is an essential first step in the understanding and conceptualization of the literature-reported factors associated with willingness to share health information. Further well-informed inquiries are possible only with this conceptualization.

Objectives

Specifically, the objectives of this review are as follows:

1. To identify peer-reviewed literature that reported factors associated with health information sharing
2. To organize factors into cohesive themes and present the synthesis as factors related to willingness to share health information

Methods

Search Strategy

A search was conducted in MEDLINE (2008-2019) to gather literature regarding willingness to share health information within the context of eHealth, which includes EHRs, personal health records, mobile health (mHealth) information, general health information, and information on social determinants of health. Additional records were also identified using Google Scholar. The search keywords included *electronic health records* AND *data sharing* OR *sharing preference* OR *willingness to share* OR *health information sharing*. The search was limited to any population that excluded health care workers or practitioners, and the participants aged ≥18 years within the US or Canadian context (see Multimedia Appendix 1 for the complete search strategy).

Identification of Records

One reviewer (IN) conducted an initial screening of the title and abstract, which identified records that were within the US and Canadian contexts and were limited to the primary peer-reviewed journal articles. Reviews, editorials, and commentaries were excluded. The screening was conducted in Excel (Microsoft Office 365; Microsoft Corp). Any ambiguous records were included for full-text screening. All included records from the title and abstract screening were saved in PDF format, and full text was screened by a single reviewer (IN). Studies were included if they reported a population aged >18 years, were not health care professionals, and reported on factors associated with sharing health information electronically or otherwise. Any ambiguity of full-text inclusion was resolved through discussion with the research team.

Synthesis

The final record listed was imported to NVivo (version 12; QSR International) for data abstraction, which was done by a single researcher (IN). This process included an aspect of thematic analysis where any factors associated with sharing health information were highlighted and coded inductively within each article. This process was carried out by a single extractor (IN). This resulted in an extensive list of factors that were distinct, overlapping, or related. Through discussion with the research team about the interrelated nature of the factors, a consensus was achieved where the factors were collated into major themes. Additional information about each record was abstracted using a predesigned Excel spreadsheet form (Microsoft Office 365; Microsoft Corporation). The extracted information included study author, publication date, study type, main objectives, population, sample size, the type of health information discussed, and major conclusions. We present a narrative synthesis on factors related to health information sharing in this report.

Results

Overview

The search was completed in October 2019. Initially, a total of 1707 records were identified through MEDLINE. Further, Google Scholar search yielded an additional 11 records for review. A total of 1650 unique records, after deduplication, were title and abstract screened, and thereafter, 1607 records were removed. Subsequently, 43 full-text articles were screened for relevance, of which 26 (60%) met the inclusion criteria and were analyzed for this review (Figure 1). The included studies in this paper reported on various populations using different methodologies. This included the general adult population using surveys, patient or hospital presenting populations (assessed using both survey and qualitative methods), and *other* groups of population, which included community-based studies or studies focusing on a particular population. The study characteristics are presented in Table 1.

Figure 1. Study flow diagram.

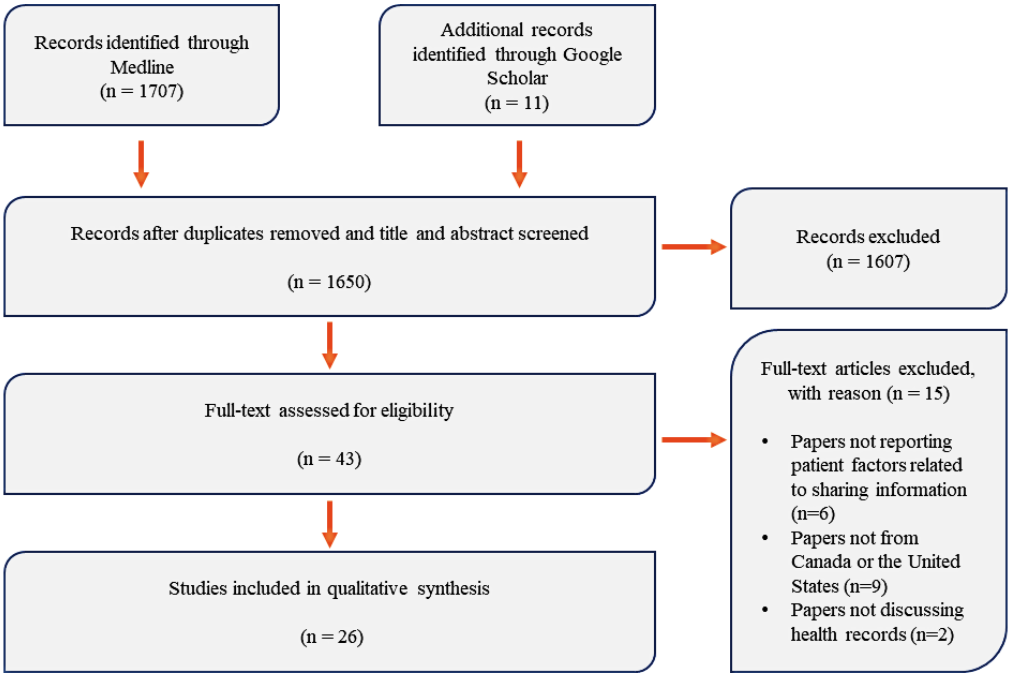


Table 1. Summary of literature reviewed.

Study	Study type	Objective	Population	Sample size, N	Health information format discussed	Factors discussed influencing health data sharing	Major findings
General adult population surveys							
Anderson and Agarwal [16]	Cross-sectional survey	Consumer willingness to provide access to patient health information to inform changes to policy	Adult public of the United States	1089	PHR[a]	• Stakeholder use of health information • Outcomes of health information • Incentives to sharing health information	Contextual factors related to the requesting stakeholder and the purpose of information being requested influence patient trust on willingness to provide health information.
Caine and Hanania [7]	Cross-sectional survey	A survey to understand patient preferences in sharing EMR[b]	Adult public receiving health care in the United States	30	EMR	• Stakeholder use of health information • Health information type and amount • Patient engagement with health information • Patient concern with data security and privacy • Patient control over data	Participants were found to have preferences in type and amount of health information shared as a function of requesting stakeholders.
Gaylin et al [17]	Cross-sectional interviews	To understand public attitudes regarding EMRs	General adult population of the United States	1014	EMR	• Income and willingness to share health information • Ethnicity and willingness to share health information • Patients concern with data privacy and security	The overall public view of using EMRs in health care delivery are positive, and participants who had previous experience with IT[c] are more likely to use and adopt EMRs.
Cocosila and Archer [18]	Cross-sectional survey	To understand the consumer motivations to implement the used of PHRs by understanding individual barriers and motivators	Adult public in Canada	772	PHR	• Stakeholder use of health information • Mode of access to health information • Age and willingness to share health information • Engagement with IT and interest in PHR • Patient engagement with health information • Patient concern with data privacy and security	Participants with and without major illness are more likely to adopt and share electronic PHRs if they perceive it as useful and an advantage to themselves. Perceptions of data security, privacy, and trust are also important.
Hasnain-Wynia et al [19]	Qualitative	To understand health plan members perceptions of the collection race, ethnicity, and primary language data	Health plan members in the United States	54	Health-related information	• Ethnicity and willingness to share health information • Patient engagement with health information • Outcome of health information	Virtually no participants in the study had problem with discussing primary language, but participants had issues with sharing information regarding their ethnicity and race.

Study	Study type	Objective	Population	Sample size, N	Health information format discussed	Factors discussed influencing health data sharing	Major findings
Donovan-Kicken et al [20]	Cross-sectional survey	To explore factors related to health literacy in the comprehension and assessment of medical disclosure and consent forms	General adult population of the United States	254	Health-related information	• Type and amount of health information shared	Health literacy and the comprehensible nature of consent documents for health research affect participation, especially with participant engagement with medical disclosure and consent documents.
Kim et al [12]	Cross-sectional survey	To understand consumer characteristics, attitudes, and beliefs regarding consent to sharing eHealth data for health care and research purposes	General adult population of the United States	800	EHR[d]	• Type and amount of health information shared • Stakeholder use of health information • Patient trust in researchers • Health information for research • Age and willingness to share health information • Patient understanding of how data are used • Patient control over data • Patient concern with data security and privacy	Individual experiences and attitudes toward sharing of EHRs needs to be considered when using EHRs for research.
Pickard and Swan [21]	Cross-sectional survey	To explore consumer attitudes toward sharing health information for research purposes	General population of the United States	128	Health-related information	• Health information type and amount • Stakeholder use of health information • Patient understanding of how data are used • Health data and management of disease • Patient engagement with other patients • Encouragement to share by stakeholders • Patient control over data • Patient concern with data security and privacy	Authors propose that health information sharing can be increased with trust, motivation, community, and informed consent.
Medford-Davis et al [13]	Cross-sectional survey	To understand patient acceptability and benefit to sharing, consent to sharing, and benefit of health records	General population of the United States	1017	EHR	• Health information type and amount • Health information for research • Patient understanding of how data are used • Outcome of health information use • Patient control over data • Patient concern with security and privacy	Most participants of the study are in favor of HIE[e] but would like more control of their health information through consent. Primary concerns with sharing health information includes concerns with privacy and security.
Spooner et al [22]	Cross-sectional survey		General adult population of the United States	3677	Health-related information	• Mode of access to health information • Stakeholder use of health information • Age and willingness to share health information	Participants of this study have high interest but low prevalence of HIE electronically.

Study	Study type	Objective	Population	Sample size, N	Health information format discussed	Factors discussed influencing health data sharing	Major findings
		To describe web-based health seeking behaviors and to identify patient-level factors to sharing of health information electronically with health care providers					
Weitzman et al [23]	Cross-sectional survey+qualitative	To investigate the willingness to share information contained in an EHR for use in public health monitoring and research	General population of the United States	181	EHR	• Type and amount of health information shared • Health information for research • Patient understanding of how data are used • Patient control over data • Patient concern with data privacy and security	High levels of willingness were found in participants in sharing EHRs with public health for the purposes of disease monitoring, evaluation, and needs assessment, as guided by themes of altruism and pragmatism.
Patient population or hospital presenting population—survey							

Study	Study type	Objective	Population	Sample size, N	Health information format discussed	Factors discussed influencing health data sharing	Major findings
Bartlett et al [24]	Cross-sectional survey	To determine the factors that impact family medicine patients' decision to allow their eHealth data to be used for research purposes	Attendees of family medicine clinics in Canada	474	eHealth data	• Age and willingness to share health information • Health information for research	Patients in family medicine clinics are more likely to refuse to contribute their deidentified eHealth data for research purposes. Relevance of the research to the patient was an impacting factor.
Brown et al [25]	Survey	A survey study to understand the desirability and functionality of a communication portal in an ICU[f]	Adult ICU patients and family in the United States	2205	eHealth data	• Stakeholder use of health information • Mode of access to health information • Age and willingness to share health information • Patient engagement with health information	Current and potential ICU patients support the feasibility and effective information sharing facilitated by an eHealth information portal. Such a portal would help in providing clinical updates, documentation of family meetings, and information regarding health care staff roles.
Garrido et al [26]	Retrospective observational study	To investigate the impact of race and ethnicity on PHR registration along with other factors	Adult members of health care network in the United States	1,764,121	PHR	• Ethnicity and willingness to share health information	Racial groups of color were less likely to register for PHRs when controlling for other factors.
Gerber et al [27]	Retrospective observational study	To understand the prevalence and patterns of PHR within an oncology population	Patents within a cancer center who had access to a secure web-based portal with their PHR in the United States	6495	EMR	• Patient engagement with IT • Ethnicity and willingness to share health information	Oncology patients readily adopt the use of EMRs. Explanatory factors are the greater health care need by these patients leads to increased portal use.
Kerath et al [28]	Cross-sectional survey	To understand attitudes related to the collection, storing, and consent toward use of genetic information for research purposes	Long Island health system patients and their families	1041	Genomic data	• Health information type and amount • Patient trust in research • Stakeholder requesting health information • Age and willingness to share health information • Patient understanding of how data are used • Previous interaction with IT • Patient concern with data security and privacy	Most participants were willing to share health information, where limitations to sharing were related to data privacy and consent procedures, along with importance of the studies being conducted.
Padrez et al [29]	Cross-sectional survey	To explore the feasibility and data availability to linking patient's social media content with their EMR data	Adult Facebook or Twitter users who presented to an emergency department	1433	EMR	• Previous engagement with IT	

Study	Study type	Objective	Population	Sample size, N	Health information format discussed	Factors discussed influencing health data sharing	Major findings
							Most individuals presenting to an emergency department that used social media consented to sharing and providing access to integrated information of their social media and EMR. The study presents a discussion on possible data repositories that link cross-platform data.
Patel et al [30]	Cross-sectional survey	To explore consumer attitudes and support for physician use of HIE within a low-income, ethnically diverse community	Adult population presenting to an emergency and ambulatory care sites	214	PHR	• Type and amount of health information shared • Stakeholder use of health information • Health information for research • Age and willingness to share health information • Health data and disease management • Outcomes of health information use • Patient concern with data security and privacy • Patient control over health data	Over half of participants supported use of PHRs by themselves and their health care providers. Potential benefits of health information influences sharing.
Pedersen et al [31]	Cross-sectional survey	To understand the acceptability of EHRs in an STI[g] clinic and its impact on intention to be screened for STI	Patients of an STI clinic in Canada	1004	EHR	• Type and amount of health information • Stakeholder use of health information • Age and willingness to share health information • Patient concern with data security and privacy	One-third of participants reported that they were not comfortable with sharing their health information and are less likely to use STI clinic.
Seltzer et al [8]	Cross-sectional survey	To explore participants willingness to share data, understand data content, and preferences related to sharing that data	Adult population presenting to an emergency department in the United States	206	Health-related information	• Type and amount of health information shared • Health information for research • Patient understanding of how data are used • Patient concern with security and privacy	Participants of the survey use a variety of modalities to generate data. Willingness to share health information for research increases for health-related insights.
Teixeira et al [32]	Cross-sectional survey	To explore attitudes of patients with HIV about having their personal health information stored and shared electronically and what factors influence their willingness to share	Patients presenting to an HIV clinic in the United States	93	PHR	• Stakeholder use of health information • Ethnicity and willingness to share health information	Results indicate patients have a high trust in their primary care provider and HIV care teams and are willing to share information with these persons.

Study	Study type	Objective	Population	Sample size, N	Health information format discussed	Factors discussed influencing health data sharing	Major findings
Weitzman et al [33]	Cross-sectional survey	To investigate attitudes and practices related to sharing health information from an EHR to support patient care and public health monitoring	Patients or guardians who used EHRs in a hospital patient portal system	261	EHR	• Type and amount of health information shared • Stakeholder use of health information • Age and willingness to share health information • Patient understanding of how data are used • Interest in PHRs • Patient engagement with health information • Patient control over data	The study found moderate levels of willingness to share electronically stored health information. Participants are more likely to share with public health authorities than are other stakeholders.
Patient population or hospital presenting population—qualitative							
Courtney (2008) [34]	Qualitative	To understand concerns regarding willingness to adopt smartphone IT in senior citizens	Adults aged ≥65 years in residential care facilities in the United States	14	Smartphone IT information collection	• Age and willingness to share health information • Engagement with other information sharers or patients • Patient concern with data privacy and security	Senior participants of this study indicate privacy as a barrier to the adoption of smartphone IT within their homes; however, their perceptions of the usefulness of the technology may be a mitigating factor.
Fuji et al (2015) [35]	Qualitative	To understand the barriers and facilitators to sustained use of PHR in patients with type 2 diabetes in managing their disease	Adult patients with type 2 diabetes in the United States	59	PHR	• Health data and management of disease • Health data and management of disease • Patient concern with data privacy and security • Patient control over health information	Patients with type 2 diabetes experience multiple benefits of using PHRs, including disease management and facilitation of behavioral change. Sustained PHR use can be achieved via building strong patient-provider relationships.
Other populations							

Study	Study type	Objective	Population	Sample size, N	Health information format discussed	Factors discussed influencing health data sharing	Major findings
Beyer, et al [36]	Observational study	To explore the implications of having community engagement in the exploring and interpretation of a GIS[h] disease mapping methodology for cancer	Rural community in the United States	60	GIS	• Community engagement with health information • Patient concern with data security	This study found that community interaction with GIS data for cancer was informative and allowed participants to build hypotheses and understanding of community health facilitating the ownership of their health data.
Jamal et al [14]	Qualitative	To understand research participant attitudes toward confidentiality and data sharing of genomic information for research purposes	Adults who consented to genomic sequencing projects in the United States	30	Genomic data	• Patient trust in researchers • Health information for research • Patient understanding of how data are used • Outcomes of health information • Patient concern with data security and privacy • Patient control over data	A complex interplay of perception of data security and privacy, individual altruism, and situational collection and use of genomic information influences information sharing.

[a]PHR: personal health record.

[b]EMR: electronic medical record.

[c]IT: information technology.

[d]EHR: electronic health record.

[e]HIE: health information exchange.

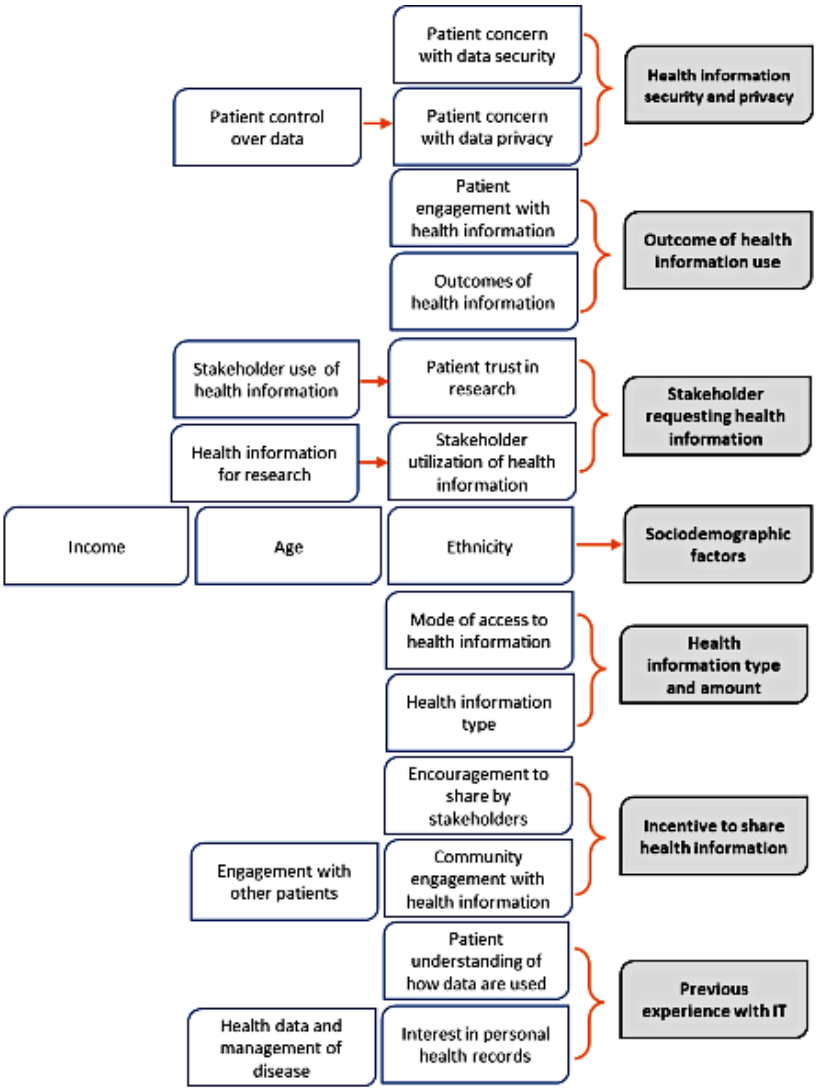
[f]ICU: intensive care unit.

[g]STI: sexually transmitted disease.

[h]GIS: geographic information system.

The inductive thematic coding process revealed multiple factors related to willingness to share health information, as reported by the study participants (Figure 2). A single study often reported a multitude of factors related to sharing information (Table 1). The factors were collated into major themes related to sharing health information. For example, multiple studies reporting an association between age, income, or ethnicity, and the willingness to share health information were grouped under the major theme *sociodemographic factors*. A similar process was followed for the remaining factors coded within the research articles, from which 7 major themes emerged. The following is a narrative synthesis of all major themes discovered in the review process.

Figure 2. Factors related to willingness to share health information inductively coded within included papers and collated into major themes. IT: information technology.



Sociodemographic Factors

A total of 15 articles reported sociodemographic factors associated with willingness to share health information. The demographic factors as a major theme were primarily noted in the survey studies, in both the general adult population and a patient or hospital presenting population. The evidence suggests an incomplete exploration of the sociodemographic factors that operate in an interrelated manner to influence the willingness to share health information. For example, the relationship between age and willingness to share health information was contested, as some studies reported that older people were more comfortable with sharing health information because of a higher level of involvement with the health care system [25], whereas other studies found them to be less willing [18] as older users were less comfortable with information technology (IT) and, therefore, less likely to share health information, this being especially true via mobile phone apps. Others have found no influence of age (or other demographic variables) to be related to sharing health information for research purposes or to improve clinical care [22,28].

Measures of social capital have an unclear association with willingness to share health information also. Some papers have reported that higher education and income increase willingness to share health information and such individuals see the benefits of sharing information [22,24,33,37], whereas others have found no influence of these factors on sharing health information [12,28,31]. It should be noted that income and education are often covariates and their individual effects on outcomes are difficult to discern. Further, mediators, such as inequitable access to technology by lower socioeconomic groups, cannot be ignored when understanding willingness to share health data [23]. Further, although ethnic disparities have been noted regarding health information sharing [26], other researchers have found no effects of ethnicity and sharing health information [12,28,31].

Incentive to Share Health Information

A total of 4 studies report the importance of incentives to increase willingness to share health information. Incentive in this case can be defined as something that drives and motivates individuals to perform an action. For the purpose of this analysis, the authors have considered incentives to be extrinsic (ie, financial) and intrinsic motivators. This theme was primarily

presented in studies sampling from the general adult population and other populations, including a community [36] and individuals who consented to share their genomic data [14]. Individuals reported various incentives that may motivate them to share health information, including monetary and material incentives such as shopping credits or money [21].

Self-management of health as a result of health information sharing is another motivator of health information sharing, including the improvements in the understanding of the participants' own health [37]. As individuals have an increased awareness of the ability to manage their health, this motivates people to share their health information. Health management can include actionable things such as knowing the likelihood of developing certain diseases, the current state of the person's health, how health affects the social environments of the person, and receiving recommendations to improve health [8,21]. For example, the day-to-day management of health markers that some mHealth apps may offer (eg, physical activity tracking, blood pressure readings, and blood glucose readings) may be an incentive for users to be more engaged with the collection and sharing of health information [35]. Further, participants may be motivated to share health information if they could connect with other individuals who shared the same health conditions [21]. This is especially relevant to mHealth apps that offer engagement with a web-based community of users.

Finally, the reviewed studies also suggested that the public fundamentally cares about the purpose for which their information is being used and is more likely to share the information if it is being used for a good purpose [29]. Participants who perceived the outcomes and implications of their health information as useful were more likely to share their health information [14].

Previous Experience With IT

A total of 12 research articles reported previous experience with IT as a factor associated with willingness to share health information. This theme was referenced by studies sampling from all types of populations, but especially so for survey studies assessing both the general and patient or hospital presenting populations. Respondents who showed interest and engagement in IT were more accepting of sharing health records [18,37]. Further, apprehension and anxiety perceived to using computers or wearables technology owing to lack of experience is a determinant of intention to share (eg, computer anxiety). Researchers argued for improving internet access and computer literacy as critical to increasing engagement and willingness to share health information, especially in a diverse population [37].

Type and Amount of Health Information

A total of 14 studies reported factors related to the type and amount of health information in association with willingness to share health information. This theme was largely reported by survey studies, of both the general population and the patient or hospital presenting population. The results suggest that individuals prefer control over the type and amount of health data, where they can control the information being shared, with a primary concern being the confidentiality of their sensitive information [31]. This may include differing sharing practices based on the sensitivity of the information being shared (eg, sexual activity or orientation, adoptions, abortions, and substance abuse) [7]. The authors found that although most participants agreed with sharing their health information, they were less likely to be tested if participants knew that their clinical information was being shared by provincial health care systems.

Data Privacy and Security

A total of 14 studies reported data privacy and security as a factor related to willingness to share health information. This theme was reported by all types of populations assessed. Within the growing trend of IT and the creation of large data repositories, security and privacy are a major concern for data producers and are closely linked to the confidentiality of sensitive information, as discussed in the previous section. Courtney [34] offers a multidimensional look into what privacy and security means within the health data field and found that patient mistrust results in withholding of health information. Fuji et al [35] found that privacy existed at both personal and technical levels, where some participants expressed themselves to be private and disliked sharing any information, whereas others stated that some technologies (eg, cloud sharing technology) may not be equipped to ensure total data security. Similar results in patients' sensitivities to sharing health information have been found in genomics research [14,28].

In practice, although health information privacy and security are valued concepts for patients when sharing their EHRs, concerns about privacy decreased in specific patient groups, such as those who were chronically ill. In such cases, the benefits of sharing medical records may have outweighed privacy risks perceptions [18]. However, Gaylin et al [17] discussed the opposite, where privacy concerns were more important than sharing health information and its potential benefits to society. Further, mitigation of privacy concerns may increase willingness to share, such as anonymization [11,23]. However, researchers discussed that with the increases in IT systems to share information (eg, using social media), individuals may still be willing to share information regardless of privacy and security concerns.

Stakeholder Requesting Health Information

Willingness to share health information is also influenced by who will use the information, which was as reported by 17 studies. This theme was primarily reported by survey studies (both general and patient or hospital presenting population). Studies showed that participants were more likely to share health information with their primary physicians, depending on the nature of the information [7]. Researchers and public organizations (nonclinical staff) were least likely to be on the list of participants' willingness to share health information [7,11,32]. Hesitancy to share was especially true when the recipients of health information were doing research that was not relevant to the participants sharing information [24]. Participants were more likely to contribute information for research purposes if they knew that it would benefit themselves or the public in some way [14].

Outcome of Health Information Use

A total of 10 studies reported that participants were influenced by the intended use and outcomes of their information when sharing health data. Again, this theme is mostly reported by survey studies, of both the general adult population and the patient or hospital presenting population. Anderson and Agarwal [16] found that the outcome and the role their health information had to play was important for sharing health information, as established trust was an important determinant of information sharing. Hasnain-Wynia et al [19] found that 90% of their study participants needed to know who was using their health information and for what purpose. Patel et al [37] found that individuals who perceived the positive benefits of sharing health information such as EHRs, such as understanding of their health, control over their health care, ability to make decisions together with their health care team, improvement in the quality of care, and satisfaction with health care, were more motivated to share their information. Brown et al [25] found that individuals who feel like they are contributing to an improvement of health care are more likely to share health information.

Discussion

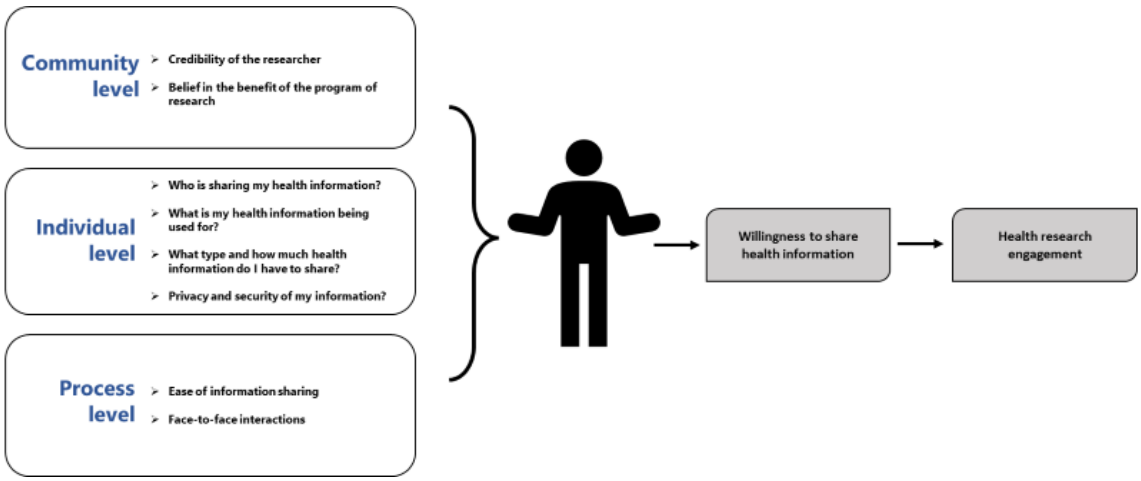
Principal Findings

Overview

The purpose of this rapid review was to locate literature that reported factors related to willingness to share health information and synthesize them into cohesive themes. Through the review process, a total of 7 major themes were discovered that explored the different aspects of the process of sharing information. This included sociodemographic factors, contextual factors (eg, type and amount of health information shared), and a mix of contextual and cognitive factors that influence willingness to share (eg, stakeholder requesting information).

The factors associated with willingness to share health information reported here ultimately suggest the importance of developing trust. Trust is complicated, and often a philosophical concept, but is generally defined as imparting authority to another and accepting the vulnerability associated with that, given that a set of expectations are met [38]. When sharing their health information, an agreement of trust is made between the individual sharing and the stakeholder accepting the information. Participants share their information accepting that they have become vulnerable by sharing their intellectual capital and personal nature of health information, and rightly expect the outcome of that sharing process to meet their expectations. It is then up to the stakeholder to upkeep those expectations, or not, ultimately building or eroding that trust. Trust during the sharing process is multifaceted, and the factors associated with willingness to share health information that were found through this study illuminate some of these facets. When assessing the overlapping and unique themes found in this study, trust seems to operate at multiple levels: (1) community level, (2) individual level, and (3) process level (Figure 3).

Figure 3. Dimensions of trust during the sharing process.



Community-Level Trust

The first is community-level trust, which speaks to themes regarding the stakeholder requesting the information and the outcomes of information shared by participants in the studies reviewed within this report. Credibility in the institutions that back up the stakeholder is important, which was especially true if the institutions were well known and had a good reputation. The credibility aspect is particularly important for certain communities who have had historically less access to power and privilege and have been exploited in the name of health research. Credibility of the stakeholders can also mean that stakeholder appreciate the diversity within communities and are willing to engage with the community to understand their perspectives [39]. Further, the relatively less willingness to share health information that is sensitive in nature may be a universal aspect of sharing for all participants, but the compounding of historical research practices, mutual stereotypes, and differences in cultures and ethnicity can influence trust building between researchers and different types of communities.

Having knowledge about the purpose, benefits, and downsides to sharing their health information was also an important factor associated with willingness to share information. Understanding that sharing health information can benefit the participants individually or benefit the entire community builds resilience and contributes to the sense of community. A recent scoping

review of barriers and facilitators of recruitment of South Asian participants found that engagement with health research was low in this population because of lack of knowledge about the scientific importance of the work, poor understanding of the research intentions, and the perception that the research benefits will not extend to their community [39].

Individual-Level Trust

The results of this study show that individual-level trust is built by a data-sharing environment where participants feel safe in sharing their health information. These factors evidently constitute a major decision-making aspect for participants when sharing health information. These factors include (1) stakeholder requesting information, (2) outcomes of health information, (3) security and privacy of health information, and (4) type and amount of health information shared. More importantly, the relative importance of these themes in this study may be because of their interrelatedness and connection with building individual-level trust through good research ethics.

The concept of data security and privacy of health data are well explored within the domain of health care, as health information is at times the most intimate, personal, and sensitive information that is maintained by the individual. Within most jurisdictions, privacy laws allow for total control over health information to the individuals, only to be disclosed if consent is authorized. Confidentiality goes a step beyond that and is usually characterized by an agreement between the individuals and the stakeholder requesting the information [40]. Indeed, participants felt that they would be more willing to share their health information if the information was going to be protected and private to a degree that they were comfortable with. Other studies have also found the sharing of information to be enhanced within the context of EHRs when privacy and security concerns were addressed [40,41].

Privacy and security of the data are closely linked to the outcome of that data, the stakeholder requesting the health information, and the type and amount of information shared. Participants are more likely to share their information if they feel they can have granular control over their shared data, which is also a form of maintaining privacy. If participants are able to control how much of their data and what type they are able to share, they have more control and feel safer in the sharing process. Participants also feel safe when they know the information is being used for its intended purpose, which is also communicated to them. For example, studies have shown the sociocultural aspects of collecting genetic information, which can be harmful or beneficial to the participants based on their familial and social circumstances [42].

Finally, who is using the health information is an important aspect of trust. Participants within the studies reviewed regularly stated they much preferred sharing their information with their physician or whomever primarily cared for them, health-wise. Studies have reported that individuals who regularly visit their physicians have a psychosocial expectation of benefit and trust from the physicians [43]. Having that interpersonal relationship built on the basis of day-to-day trust may be an important aspect to creating a space where health information sharing can occur. The lack of sharing of participants to other stakeholders, including organizations not associated with the health care of the participant, also points to the lack of trust and skepticism about the maintenance of privacy by these organizations.

Process-Level Trust

There is a paucity of literature describing the process of information sharing as having a role in participants' willingness to share their information. Within populations, the ease of the information-sharing process can have a large influence on whether or not a participant will engage in sharing their health information. Factors as simple as language barriers, health literacy, and type of data collection instrument can determine a study's success in engaging its population of interest [44]. In addition, complex factors, such as the sociodemographic diversity within a community, must also be addressed. For example, some ethnocultural communities may have first- and second-generation migrants who may have differing needs when it comes to ease of information sharing, where older-generation participants may require translation services or a different mode of data collection (face-to-face vs on the phone) to successfully share their information [45].

Implications of Findings

The results of this study suggest actionable items that stakeholders can consider, including introducing policy changes that aim to develop a mutually beneficial information-sharing partnership between the communities of interest. Further, in order to motivate individuals to share their health information, their situations within their community must be appreciated, and equal power should be divided among the researcher and community members on the control and direction of the data-sharing partnership [46]. This is compared with researchers controlling the collection, analysis, and dissemination of the information along with reaping its benefits, with little input from participants. To build effective data-sharing partnerships, researchers should be able to work in collaboration with community members and understand the community living, working, and socializing conditions. To do this, credible and respectful access to the community should be pursued by building relationships with community champions and organizations that have a long-standing dedication to their communities. As suggested by the findings of this review, this can be done through training and development of guidelines that assist within building such relationships, which can exist at the institutional and national research level.

Another suggested actionable item could be the documentation of the process of rapport and building relationships with communities regarding building an information-sharing partnership, along with a systematic way of collecting the community perspectives on barriers and facilitators to sharing information. Although there is a great amount of literature using and describing methodologies that view research participants and the community as partners throughout the research process (eg, community-based participatory research and integrated knowledge translation), more exploration is required to create policies and guidelines for effective documentation of information-sharing partnerships.

A deeper understanding of conducting ethical research, the abstract nature of maintaining confidentiality, and respect for the individuals and their experiences is essential throughout the information-sharing process to develop trust. For example, many research studies suffer from the simplistic assumption that a single consent form is enough to assure ethical standards for their participants. However, the results of this study show that, within a community, more is needed. Indeed, a study can maintain excellent privacy and confidentiality within their protocol but may still conduct research that is framed in a way that is disrespectful toward certain ethnocultural communities [47]. Therefore, a reassessment of research ethics evaluation processes at the institutional levels may need to be improved and adjusted to address differences in conducting research in data-sharing communities.

Sharing of health information that is easy, accessible, and feasible for the participant can also cultivate trust. Having evidence-informed standards and clear guidelines for collecting health information can not only benefit stakeholders interested in gaining information by increasing reproducibility but also benefit the information-sharing partnership [48]. That being said, stakeholders should consider the population they are hoping to collect information from when choosing or creating these standards. For example, simply measuring the concept of ethnicity in populations can be difficult, as some participants may not see their ethnicity, or diversity within an ethnicity, being reflected in the type of questionnaire they are given. Further incentives are known to increase research engagement and may be an important aspect of building information-sharing partnerships in ethnocultural communities. However, simple financial incentives may not be enough to garner continued information sharing, but rather, more customized incentives may be needed for the communities that researchers are interested in. Studies have demonstrated that incentives for ethnic and minority communities, such as colearning activities and a chance to contribute to the research development, are sustainable incentives that build trusting partnerships [31].

Limitations, Strengths, and Next Steps

This review is limited by its rapid review methodology, which fails to conduct a broader search of the literature and critically analyze the included studies. For example, this review contains studies with a variable sample size, which could influence the generalizability of the results of certain studies with smaller sample sizes. Further, the included studies report on the incomparable context of individuals, where some participants are hospitalized patients, as compared with the general adult population (Table 1). When assessing the results of this study, there are some notable differences in results when comparing the population assessed and the methods used to assess them. For example, some themes are overly represented in survey studies in both the general adult population and a patient or hospital presenting population simply because of the methods used. In survey studies, authors can quickly measure study participant preferences on the type and amount of health information shared, outcomes of health information use, and the stakeholder requesting health information. Further, most survey studies feature a larger population size, which can also influence the results by the inclusion of more viewpoints and more possible factors that influence willingness to share. However, the study finds its strengths in the reporting of concise narrative synthesis of factors associated with willingness to share health information into cohesive themes and subsequent domains, using thematic coding methods. An important next step for this review would be a systematic search of the literature, allowing for an in-depth analysis of health information sharing. Further, primary studies focusing on health information exchange in populations facing health disparities are warranted to expand the field.

Conclusions

This review provided a concise report on factors associated with willingness to share health information, including a conceptual framework that outlined sociodemographic, cognitive, and contextual domains associated with health information sharing. Further, this review emphasized the importance of data generators' viewpoints and the complex systems of factors that shape their decision to share health information. The factors related to information sharing reported in this review have important implications in participant engagement and reorientation of methodologies in research studies to build sustained information exchange capacity. Sustained information exchange is an important aspect of current trends in medical research and public health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
MEDLINE search strategy.
[DOCX File , 25 KB - humanfactors_v9i1e20702_app1.docx]

References

1. Prosperi M, Min JS, Bian J, Modave F. Big data hurdles in precision medicine and precision public health. BMC Med Inform Decis Mak 2018 Dec 29;18(1):139 [FREE Full text] [doi: 10.1186/s12911-018-0719-2] [Medline: 30594159]
2. van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst AJ, et al. A systematic review of barriers to data sharing in public health. BMC Public Health 2014 Nov 05;14(1):1144 [FREE Full text] [doi: 10.1186/1471-2458-14-1144] [Medline: 25377061]

3. Morgan A, Mooney S, Aronow B, Brenner S. Precision medicine: data and discovery for improved health and therapy. In: Proceedings of the Pacific Symposium - Biocomputing 2016. 2016 Presented at: Pacific Symposium - Biocomputing 2016; January 4–8, 2016; Kohala Coast, Hawaii, USA p. 243-248. [doi: 10.1142/9789814749411_0023]
4. Leff DR, Yang G. Big data for precision medicine. Engineering 2015 Sep;1(3):277-279. [doi: 10.15302/j-eng-2015075]
5. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. Am J Prev Med 2016 Mar;50(3):398-401 [FREE Full text] [doi: 10.1016/j.amepre.2015.08.031] [Medline: 26547538]
6. Zakim D, Schwab M. Data collection as a barrier to personalized medicine. Trends Pharmacol Sci 2015 Feb;36(2):68-71. [doi: 10.1016/j.tips.2014.11.002] [Medline: 25479798]
7. Caine K, Hanania R. Patients want granular privacy control over health information in electronic medical records. J Am Med Inform Assoc 2013 Jan 01;20(1):7-15 [FREE Full text] [doi: 10.1136/amiajnl-2012-001023] [Medline: 23184192]
8. Seltzer E, Goldshear J, Guntuku SC, Grande D, Asch DA, Klinger EV, et al. Patients' willingness to share digital health and non-health data for research: a cross-sectional study. BMC Med Inform Decis Mak 2019 Aug 08;19(1):157 [FREE Full text] [doi: 10.1186/s12911-019-0886-9] [Medline: 31395102]
9. Lodhia V, Karanja S, Lees S, Bastawrous A. Acceptability, usability, and views on deployment of peek, a mobile phone mHealth intervention for eye care in Kenya: qualitative study. JMIR Mhealth Uhealth 2016 May 09;4(2):e30 [FREE Full text] [doi: 10.2196/mhealth.4746] [Medline: 27160779]
10. Madden TJ, Ellen PS, Ajzen I. A comparison of the theory of planned behavior and the theory of reasoned action. Pers Soc Psychol Bull 2016 Jul 02;18(1):3-9. [doi: 10.1177/0146167292181001]
11. Whiddett R, Hunter I, Engelbrecht J, Handy J. Patients' attitudes towards sharing their health information. Int J Med Inform 2006 Jul;75(7):530-541. [doi: 10.1016/j.ijmedinf.2005.08.009] [Medline: 16198142]
12. Kim KK, Sankar P, Wilson MD, Haynes SC. Factors affecting willingness to share electronic health data among California consumers. BMC Med Ethics 2017 Apr 04;18(1):25 [FREE Full text] [doi: 10.1186/s12910-017-0185-x] [Medline: 28376801]
13. Medford-Davis LN, Chang L, Rhodes KV. Health Information Exchange: what do patients want? Health Informatics J 2017 Dec 31;23(4):268-278 [FREE Full text] [doi: 10.1177/1460458216647190] [Medline: 27245671]
14. Jamal L, Sapp JC, Lewis K, Yanes T, Facio FM, Biesecker LG, et al. Research participants' attitudes towards the confidentiality of genomic sequence information. Eur J Hum Genet 2014 Aug 27;22(8):964-968 [FREE Full text] [doi: 10.1038/ejhg.2013.276] [Medline: 24281371]
15. Tricco AC, Antony J, Zarin W, Strifler L, Ghassemi M, Ivory J, et al. A scoping review of rapid review methods. BMC Med 2015 Sep 16;13(1):224 [FREE Full text] [doi: 10.1186/s12916-015-0465-6] [Medline: 26377409]
16. Anderson CL, Agarwal R. The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information. Inform Syst Res 2011 Sep;22(3):469-490. [doi: 10.1287/isre.1100.0335]
17. Gaylin D, Moiduddin A, Mohamoud S, Lundeen K, Kelly J. Public attitudes about health information technology, and its relationship to health care quality, costs, and privacy. Health Serv Res 2011 Jun;46(3):920-938 [FREE Full text] [doi: 10.1111/j.1475-6773.2010.01233.x] [Medline: 21275986]
18. Cocosila M, Archer N. Perceptions of chronically ill and healthy consumers about electronic personal health records: a comparative empirical investigation. BMJ Open 2014 Jul 23;4(7):e005304 [FREE Full text] [doi: 10.1136/bmjopen-2014-005304] [Medline: 25056975]
19. Hasnain-Wynia R, Taylor-Clark K, Anise A. Collecting race, ethnicity, and language data to identify and reduce health disparities: perceptions of health plan enrollees. Med Care Res Rev 2011 Jun 01;68(3):367-381. [doi: 10.1177/1077558710382549] [Medline: 20889576]
20. Donovan-Kicken E, Mackert M, Guinn TD, Tollison AC, Breckinridge B, Pont SJ. Health literacy, self-efficacy, and patients' assessment of medical disclosure and consent documentation. Health Commun 2012 Aug;27(6):581-590. [doi: 10.1080/10410236.2011.618434] [Medline: 22107084]
21. Pickard K, Swan M. Big desire to share big health data: a shift in consumer attitudes toward personal health information. Semnatic Scholar. 2014. URL: https://tinyurl.com/2p8wv9xr [accessed 2022-01-24]
22. Spooner K, Salemi J, Salihu H, Zoorob R. eHealth patient-provider communication in the United States: interest, inequalities, and predictors. J Am Med Inform Assoc 2017 Apr 01;24(e1):18-27 [FREE Full text] [doi: 10.1093/jamia/ocw087] [Medline: 27497797]
23. Weitzman ER, Kaci L, Mandl KD. Sharing medical data for health research: the early personal health record experience. J Med Internet Res 2010 May 25;12(2):e14 [FREE Full text] [doi: 10.2196/jmir.1356] [Medline: 20501431]
24. Bartlett G, Macgibbon B, Rubinowicz A, Nease C, Dawes M, Tamblyn R. The importance of relevance: willingness to share eHealth data for family medicine research. Front Public Health 2018 Sep 4;6:255 [FREE Full text] [doi: 10.3389/fpubh.2018.00255] [Medline: 30234095]
25. Brown SM, Bell SK, Roche SD, Dente E, Mueller A, Kim T, et al. Preferences of current and potential patients and family members regarding implementation of electronic communication portals in intensive care units. Annals ATS 2016 Mar;13(3):391-400. [doi: 10.1513/annalsats.201509-638oc]
26. Garrido T, Kanter M, Meng D, Turley M, Wang J, Sue V, et al. Race/ethnicity, personal health record access, and quality of care. Am J Manag Care 2015 Feb 01;21(2):103-113 [FREE Full text] [Medline: 25880485]

27. Gerber DE, Laccetti AL, Chen B, Yan J, Cai J, Gates S, et al. Predictors and intensity of online access to electronic medical records among patients with cancer. J Oncol Pract 2014 Sep;10(5):307-312. [doi: 10.1200/jop.2013.001347]
28. Kerath SM, Klein G, Kern M, Shapira I, Witthuhn J, Norohna N, et al. Beliefs and attitudes towards participating in genetic research - a population based cross-sectional study. BMC Public Health 2013 Feb 07;13(1):114 [FREE Full text] [doi: 10.1186/1471-2458-13-114] [Medline: 23388465]
29. Padrez KA, Ungar L, Schwartz HA, Smith RJ, Hill S, Antanavicius T, et al. Linking social media and medical record data: a study of adults presenting to an academic, urban emergency department. BMJ Qual Saf 2016 Jun 13;25(6):414-423. [doi: 10.1136/bmjqs-2015-004489] [Medline: 26464519]
30. Patel VN, Dhopeshwarkar RV, Edwards A, Barron Y, Likourezos A, Burd L, et al. Low-income, ethnically diverse consumers' perspective on health information exchange and personal health records. Inform Health Soc Care 2011 Dec 18;36(4):233-252. [doi: 10.3109/17538157.2011.554930] [Medline: 21851182]
31. Pedersen H, Taylor D, Gilbert M, Achen M, Lester R, Ogilvie G. A cross-sectional survey exploring attitudes towards provincial electronic health record implementation among clients attending the Provincial Sexually Transmitted Infections Clinic in British Columbia. Sex Transm Infect 2015 Feb 05;91(1):44-48. [doi: 10.1136/sextrans-2014-051804] [Medline: 25480149]
32. Teixeira PA, Gordon P, Camhi E, Bakken S. HIV patients' willingness to share personal health information electronically. Patient Educ Couns 2011 Aug;84(2):9-12 [FREE Full text] [doi: 10.1016/j.pec.2010.07.013] [Medline: 20724095]
33. Weitzman ER, Kelemen S, Kaci L, Mandl KD. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. BMC Med Inform Decis Mak 2012 May 22;12(1):39 [FREE Full text] [doi: 10.1186/1472-6947-12-39] [Medline: 22616619]
34. Courtney KL. Privacy and senior willingness to adopt smart home information technology in residential care facilities. Methods Inf Med 2008 Jan 19;47(1):76-81. [doi: 10.3414/me9104] [Medline: 18213432]
35. Fuji KT, Abbott AA, Galt KA. A qualitative study of how patients with type 2 diabetes use an electronic stand-alone personal health record. Telemed J E Health 2015 Apr;21(4):296-300 [FREE Full text] [doi: 10.1089/tmj.2014.0084] [Medline: 25614996]
36. Beyer KM, Comstock S, Seagren R. Disease maps as context for community mapping: a methodological approach for linking confidential health information with local geographical knowledge for community health research. J Community Health 2010 Dec 30;35(6):635-644. [doi: 10.1007/s10900-010-9254-5] [Medline: 20352481]
37. Patel VN, Abramson E, Edwards AM, Cheung MA, Dhopeshwarkar RV, Kaushal R. Consumer attitudes toward personal health records in a beacon community. Am J Manag Care 2011 Apr;17(4):104-120 [FREE Full text] [Medline: 21774099]
38. Platt J, Kardia S. Public trust in health information sharing: implications for biobanking and electronic health record systems. J Pers Med 2015 Feb 03;5(1):3-21 [FREE Full text] [doi: 10.3390/jpm5010003] [Medline: 25654300]
39. Quay TA, Frimer L, Janssen PA, Lamers Y. Barriers and facilitators to recruitment of South Asians to health research: a scoping review. BMJ Open 2017 Jun 01;7(5):e014889 [FREE Full text] [doi: 10.1136/bmjopen-2016-014889] [Medline: 28576896]
40. Gostin LO, Turek-Brezina J, Powers M, Kozloff R. Privacy and security of health information in the emerging health care system. Health Matrix Clevel 1995;5(1):1-36. [Medline: 10141742]
41. Ancker JS, Edwards AM, Miller MC, Kaushal R. Consumer perceptions of electronic health information exchange. Am J Prev Med 2012 Jul;43(1):76-80. [doi: 10.1016/j.amepre.2012.02.027] [Medline: 22704751]
42. Burgess MM. Beyond consent: ethical and social issues in genetic testing. Nat Rev Genet 2001 Feb;2(2):147-151. [doi: 10.1038/35052579] [Medline: 11253055]
43. Appari A, Johnson ME. Information security and privacy in healthcare: current state of research. Int J Internet Enterpr 2010;6(4):279. [doi: 10.1504/ijiem.2010.035624]
44. van der Velde J, Williamson DL, Ogilvie LD. Participatory action research: practical strategies for actively engaging and maintaining participation in immigrant and refugee communities. Qual Health Res 2009 Sep 18;19(9):1293-1302. [doi: 10.1177/1049732309344207] [Medline: 19690209]
45. Hennekam S, Bacouel-Jentjens S, Yang I. Perceptions of diversity management practices among first- versus second-generation migrants. Work Employ Society 2019 Nov 26;34(5):844-863. [doi: 10.1177/0950017019887335]
46. Anderson EE, Solomon S, Heitman E, DuBois JM, Fisher CB, Kost RG, et al. Research ethics education for community-engaged research: a review and research agenda. J Empir Res Hum Res Ethics 2012 Apr 01;7(2):3-19. [doi: 10.1525/jer.2012.7.2.3] [Medline: 22565579]
47. Godard B, Ozdemir V, Fortin M, Egalité N. Ethnocultural community leaders' views and perceptions on biobanks and population specific genomic research: a qualitative research study. Public Underst Sci 2010 Jul 12;19(4):469-485. [doi: 10.1177/0963662509104721] [Medline: 20977184]
48. Hajduk GK, Jamieson NE, Baker BL, Olesen OF, Lang T. It is not enough that we require data to be shared; we have to make sharing easy, feasible and accessible too!. BMJ Glob Health 2019 Jul 22;4(4):e001550 [FREE Full text] [doi: 10.1136/bmjgh-2019-001550] [Medline: 31406588]

Abbreviations

EHR: electronic health record
IT: information technology
mHealth: mobile health

Edited by A Kushniruk; submitted 26.05.20; peer-reviewed by J Jones, K Fuji; comments to author 14.07.20; revised version received 30.11.20; accepted 11.10.21; published 09.02.22.

Please cite as:
Naeem I, Quan H, Singh S, Chowdhury N, Chowdhury M, Saini V, TC T
Factors Associated With Willingness to Share Health Information: Rapid Review
JMIR Hum Factors 2022;9(1):e20702
URL: https://humanfactors.jmir.org/2022/1/e20702
doi:10.2196/20702
PMID:35138263

Review

Coaching to Support Mental Health Apps: Exploratory Narrative Review

Ashley Meyer[1], BA; Hannah Wisniewski[1], BS; John Torous[1], MBI, MD

Division of Digital Psychiatry, Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, United States

Corresponding Author:
John Torous, MBI, MD
Division of Digital Psychiatry
Department of Psychiatry
Beth Israel Deaconess Medical Center
330 Brookline Ave
Boston, MA, 02115
United States
Phone: 1 6176676700
Email: jtorous@bidmc.harvard.edu

Abstract

Background: The therapeutic alliance is crucial for the success of face-to-face therapies. Little is known about how coaching functions and fosters the therapeutic alliance in asynchronous treatment modalities such as smartphone apps.

Objective: The aim of this paper was to assess how coaching functions and fosters the therapeutic alliance in asynchronous treatment modalities.

Methods: We conducted a selected review to gather preliminary data about the role of coaching in mobile technology use for mental health care. We identified 26 trials using a 2019 review by Tønning et al and a 2021 scoping review by Tokgöz et al to assess how coaching is currently being used across different studies.

Results: Our results showed a high level of heterogeneity as studies used varying types of coaching methods but provided little information about coaching protocols and training. Coaching was feasible by clinicians and nonclinicians, scheduled and on demand, and across all technologies ranging from phone calls to social media.

Conclusions: Further research is required to better understand the effects of coaching in mobile mental health treatments, but examples offered from reviewed papers suggest several options to implement coaching today. Coaching based on replicable protocols that are verifiable for fidelity will enable the scaling of this model and a better exploration of the digital therapeutic alliance.

(JMIR Hum Factors 2022;9(1):e28301) doi:10.2196/28301

KEYWORDS

smartphone; apps; mental health; coaching; engagement

Introduction

Smartphones and other mobile technologies are increasingly used in mental health care. The COVID-19 pandemic has highlighted the need for mobile treatments in providing access to and augmenting mental health care. However, fundamental questions remain around app engagement and efficacy as do concerns about technology use in the contact of coaching support and the therapeutic alliance. The therapeutic alliance, which is described as the alliance between a clinician and a patient [1], is considered crucial to the success of face-to-face therapies and associated with successful outcomes [2]. Numerous meta-analyses have confirmed the central role of the therapeutic alliance in driving both engagement and efficacy across both face-to-face care and even telehealth video visits. However, less is known about how this alliance functions in asynchronous treatment modalities such as smartphone apps. In both research studies as well as commercial apps, a rise in coaching to support engagement may be conceptualized in the context of adding an element of the therapeutic alliance into digital care. Preliminary research suggests that coaching is feasible and acceptable, with both clinicians and clients identifying benefits to this additional support, such as increased motivation and guidance as well as a new way to focus on clinical work [3]. Yet, little is known about how this app coaching is delivered or impacts outcomes.

This lack of knowledge on the therapeutic alliance and coaching around smartphone apps is in part related to a lack of consensus on coaching methods and training. Such a lack of protocols or training manuals has immediate consequences as well-trained clinicians are more prepared to foster the therapeutic alliance and support technological difficulties [4]. At the beginning of the COVID-19 pandemic, 21% of community health centers across the United States lacked training for telehealth [5]. Training for mobile apps is nascent, and evidence can often only be found in the research literature. For example, one study in Australia found that clinician training on mobile technologies and protocols for consistent messaging with the patient was necessary to increase engagement and knowledge of the apps [6]. A recent report from Kaiser Permanente explained that teaching clinical staff how to use and interact with apps in care settings was critical for implementation success [7]. With COVID-19, grassroots efforts to train medical students to support patients around app use have also emerged [8]. The need for proper training and new knowledge around engagement and alliance to support app use is thus an important new facet of care toward offering accessible mental health care in an increasingly digital world. The term "training" is used throughout this work and defined based on the above references as one or more of the following: the receipt of coaching protocols or manuals by the coach, instruction for the coach on how to conduct telehealth sessions, how to use and interact with mobile apps, and how to support patients around app use.

Several solutions have been proposed already. Many have examined a more traditional "coaching" model in which a member of the care team maintains contact with patients between visits to foster engagement with the mobile technology. This contact may be through the app, text messages, or even phone calls. A more engaged approach is conceptualized around the concept of a "digital navigator," which seeks to integrate into more aspects of care with the goal of supporting both the patient and clinician [9]. Ben-Zeev et al [10] and Noel et al [11] have proposed related roles called clinical technology specialists or technology specialists, respectively. Each of these positions seeks to provide technology recommendations and support user engagement with digital technologies.

A related method for maintaining alliance with patients using digital technology uses a "social" model of coaching. This model focuses on interactions between peers as opposed to a member of the clinical or study team. Although this model of coaching appears less used, several studies have incorporated this method. McEnery et al [12] evaluated the feasibility of an online intervention, EMBRACE, where participants maintained contact with clinical moderators, a more traditional coach, as well as peer-to-peer moderators, who were young individuals with lived mental health experience who encouraged participant engagement and provided support. Further, Alvarez-Jimenez et al [13] assessed the feasibility and acceptability of the enhanced moderated online social therapy (MOST+), which allowed participants to interact with other participants using the platform, as well as peer moderators who facilitated engagement. However, neither of these protocols directly assessed the effect of moderation on outcomes, and, of note, neither involved smartphone apps.

To further understand the current knowledge of coaching effects, we conducted an investigation to gather preliminary data on coaching and to understand its effect on engagement and outcomes. We hypothesized that there would be substantial variability and little consensus on coaching protocols, inconsistent reporting measures, and a lack of protocols that directly assess coaching effects. However, understanding how coaching is reported and broad trends in its outcomes is useful for new efforts to align new research and implementation efforts with prior work. In turn, understanding and identifying best practices to facilitate coaching and support the digital therapeutic alliance is crucial as remote psychotherapy increases in popularity and necessity.

Methods

We conducted a selected review to gather preliminary data about the role of coaching in mobile technology use. Realizing there is no simple means to identify relevant papers as all will have some degree of coaching support, often unreported in research assistant help, we opted for an exploratory sampling approach. We chose to use a prior review of randomized controlled trials on smartphone-based treatment in psychiatry as well as a recent scoping review on digital health interventions for depression (featuring 6 smartphone-based interventions) as the samples for assessing how coaching is offered across different apps and studies. We did not attempt to conduct a qualitative analysis as we expected outcomes to be heterogeneous and diverse given the state of the literature.

We identified the trials using the 2019 review by Tønning et al [14] on the methodological challenges of randomized controlled trials on smartphone-based treatment in psychiatry and the 2021 scoping review by Tokgöz et al [15] on digital health interventions for depression. We selected these papers as they offered a recent and comprehensive sample of studies from which we could explore coaching. Each trial featured in this review was read and coded by the authors for method of contact with the patient, training coaches received, on demand vs scheduled interactions, clinical vs nonclinical interactions, evidence of dose effect, and social vs coaching model. Studies were excluded if the participants had no interaction while using the mobile technology.

After coding of each trial was completed, the trials were sorted into the following 4 categories based on the frequency (scheduled vs on demand) and nature of the coaching (nonclinical vs clinical): scheduled or clinical; scheduled or nonclinical; on demand or clinical; and on demand or nonclinical. Scheduled coaching included coaching delivered on a set time frame, such as once per week or per month, or after the completion of a certain assessment. On-demand coaching was delivered on an irregular schedule based on the needs of the study participant or clinician, such as clinician responses to high participant assessment score or a participant contacting study team for questions; however, we did not include on-demand crisis intervention in this category and will not be reviewing such interventions within this work. Clinical coaching focused on the participant's symptomatology, whereas nonclinical coaching focused on technology or study protocol

questions. Some trials that included a variety of methods for interaction were coded into multiple categories. We then assigned a hierarchy to the codes as follows from highest priority to lowest priority: (1) on demand or clinical; (2) on demand or nonclinical; (3) scheduled or clinical; and (4) scheduled or nonclinical. Dose effect was defined as a reported association between the time or intensity of coaching for the participants and a primary study outcome. While the digital therapeutic alliance is of critical interest, it is not yet possible to code given that the means to assess it are nascent, as discussed later in this paper.

Results

A total of 32 trials were reviewed as featured in the 2019 review by Tønning et al and 2021 scoping review by Tokgöz et al; 6 of these studies did not involve reported coaching interaction while the participants were using the smartphone technology and were excluded from review by the study team [16-21]. Therefore, a total of 26 trials were included [22-47]. A summary of the results can be seen in Multimedia Appendix 1 and Table 1.

As seen in Table 1, there was high variability around coaching across each of the 26 studies in the type of coaching delivered and coach training. The majority of studies included a scheduled coaching component (14 scheduled or clinical) [22,23,25,26,28,29,31,33,36,38,40,45-47]; 12 scheduled or nonclinical [24,29,32,33,35,38,41,45,46]) compared to an on-demand coaching component (8 on demand or clinical [22,24,27,29,30,37-39]; 9 on demand or nonclinical [24,25,29,34,37,38,42,44,46]); 11 of the studies incorporated 2 or more kinds of coaching [18,20,21,25,29,33,34,38,44-46].

There was less variability as to who provided the coaching. Clinicians acted as coaches in 18 of the studies [22-31,36-41,45,47], nonclinicians in 3 studies [32,35,42], and peers in 3 studies [34,37,44]. Moreover, 4 studies did not specify who acted as coaches [33,34,44,46]. In addition, the majority of studies did not specify the type of training that the coaches completed for their role; 16 studies did not specify the type of training [22-25,27,30,32,33,35-45], 1 study specified that training was not conducted [34], and 7 studies specified that coaches underwent training of some kind [26,28,29,31,38,41,45]. Of the studies that did specify the training for the coaches, there was high variability; 1 study noted the coaches' training was their standard training as a part of their clinical psychology program [29]; another noted the training was a "1-day workshop in using the self-help program and on how to write the weekly feedback, based on case material from earlier trials" [38], while another only stated their training was "based on the supportive accountability model" [41]. Only 1 study offered a protocol for the training offered [46].

Only 3 studies used the social coaching model [34,37,44]; 1 of these trials used only the social model [34], while 2 of them used the social model along with the coach model [37,44]. Boettcher et al [34], who used only the social model, examined the efficacy of a smartphone app called Challenger in reducing anxiety symptoms in individuals with social anxiety disorder. The participants were randomized to use Challenger and a self-help program simultaneously, the self-help program for 6 weeks followed by the Challenger app for 6 weeks, or a waitlist control. Challenger used cognitive behavioral therapy techniques to encourage its users to complete small exposure and behavioral challenges in everyday life. The skills gradually increased in difficulty. After each skill, the user is able to complete a reflection of the task, which is sent to another user who is able to respond with constructive feedback. The participants in Schlosser et al [37] and Roepke et al [44] were able to interact with other app users or use a forum and recruit social support from Facebook, respectively.

There was evidence of a dose effect in only 1 study. In particular, a pilot randomized controlled trial conducted by Pfeiffer et al [47], exploring psychotherapeutic text messaging for depression, found that change in behavioral activation was correlated with specifically 6 weeks of receiving acceptance and commitment therapy-based messages (ρ=-0.25; P<.05), as opposed to 12 weeks, at which point there was no correlation observed [47]. Studies used varying measures of engagement and efficacy of the respective smartphone technologies; 38.5% (n=10) reported percent completion of the program [21-23,28,29,32,34,35,40,46], 19.2% (n=5) reported app use per week or day [24,27,30,36,46], 19.2% (n=5 reported the retention or dropout rate [37,38,43,45,46], and 7.7% (n=2) reported the number of logins to the program [33,44]. However, only 10 studies reported the duration of time spent per coaching interaction [22-25,29-31,41,42,46], and many did not directly assess the influence of coaching on the results.

Finally, there was high variability in the mode of contact used across the studies; 13 (50%) of the studies used 2 or more means of contact [22,24,25,27,29,30,32,37,41,42,44-46]. Phone calls were most commonly used to contact participants (14, 53.8%) [22-24,27,29,30,32,33,35,37,41,42,45,46], followed by emails (7, 26.9%) [22,25,27,32,41,43,44], then in-person meetings (6, 23%) [26,28,31,36,42,45], in-app messaging (5, 19.2%) [24,29,30,34,37], text messaging (5, 19.2%) [25,27,38,46,47], FaceTime or teleconference (2, 7.7%) [37,39], app notifications (1, 3.8%) [39], and Facebook (1, 3.8%) [44].

Table 1. Summary of coding metrics.

Criteria and coding specifications	Number of studies, n (%)
Type of training	
Specified	9 (35)
Unspecified	16 (62)
Mode of contact	
Email	7 (27)
Phone call	14 (54)
In-app message	5 (19)
Text message	5 (19)
In-person	6 (23)
FaceTime or teleconference	2 (8)
Notifications	1 (4)
Facebook	1 (4)
On demand vs scheduled	
On demand	12 (46)
Scheduled	21 (81)
Clinical vs nonclinical	
Clinical	19 (73)
Nonclinical	14 (54)
Clinician vs nonclinician vs peer	
Clinician	18 (69)
Nonclinician	3 (12)
Peer	3 (12)
Not specified	4 (15)
Time spent per interaction	
Specified	11 (42)
Not specified	15 (58)
Evidence of dose effect	
No	25 (96)
Yes	1 (4)
Social vs coach model	
Social	24 (92)
Coach	5 (19)
Participants compensated	
Yes	17 (65)
No	9 (45)
Participants received smartphone	
Yes	4 (15)
Yes, if necessary	7 (27)
No	16 (62)
Remote study	
Yes	12 (46)
No	14 (54)

Discussion

Coaching offers a solution to engagement challenges with digital mental health, but its interpretation and implementation remain heterogeneous, consistent with our hypothesis. While our results are not a comprehensive review, they offer a selected sample across the mental health app literature, which highlights the diversity of efforts and results when applying different models of coaching to support apps. A lack of consensus around coaching protocols and outcomes precludes discussion of whether coaching may be a covariate, confounder, moderator, or mediator for clinical improvement with apps. While we were not able to explicitly measure the therapeutic alliance construct within this work, the heterogeneity found across coaching modalities may suggest a lack of consensus regarding how to best foster a digital therapeutic alliance [48] between the patient and clinicians. Recent studies not captured in our sample have employed the Digital Working Alliance Inventory to measure alliance with apps and suggested that such an alliance may predict app engagement [49], highlighting the significance of future research and standardization around this concept.

The high degree of heterogeneity reflected in our results suggests the versatility of coaching and its ability to easily adapt to unique circumstances. Coaching was feasible across all platforms ranging from text messages to social media and for both on-demand and scheduled interactions. Coaches were also able to support completely remote studies (defined as specifically involving no synchronous interactions) as well as offer face-to-face services in meeting with participants in other studies. While a clinician served the role in 69.2% of studies, the role is also accessible to other people including those with no formal training.

One challenge around understanding the efficacy of coaching, beyond the heterogeneity of the role and studies, is that training protocols, fidelity to those protocols, and coaching specific outcomes are often not reported. Without understanding how coaches are trained and if they adhere to that training during the study, it is impossible to understand what support is actually being delivered. Study metrics reports by coach instead of participant and cohort may also offer productive data toward understanding the impact of this role. While no studies measured outcomes such as the Working Alliance Inventory, alliance-specific measures would offer information into potential mechanisms of action.

However, the results from this paper offer several paths forward. These results suggest that clinical vs nonclinical staff can serve in coaching roles, and scheduled vs on-demand support can also both be feasible. Crowdsourcing peer support via social networks or small internal networks also appears feasible. As the role and best practices evolve, clinics can implement the methods that best match their local needs and resources. The different models presented in this paper can serve as examples in building new coaching services and provide measures to consider during implementation. While beyond the immediate scope of this article, protocols around digital mental health coaching are emerging and can serve as further reference [50,51]. Of note, neither of these protocols or earlier versions of them was used in any paper reviewed.

Our results are in line with prior works that have examined coaching around mental health apps. In a 2020 paper, Callejas et al [51] reported on selected examples and noted a need for more data around engagement and mechanisms of action underlying coaching. A lack of consensus around app engagement measures has also been found in recent reviews [52,53].

A chief limitation of this work is that it draws a sample from only 2 reviews of mental health app studies. Given that nearly every digital health study involves some degree of coaching (even if they are informal support from research assistants, which may not be reported), it is infeasible to conduct a broader review. Therefore, our goal was not to include every relevant paper, but rather to conduct a preliminary investigation into coaching techniques used by recent studies and identify trends. Other studies have specifically explored coaching and mental health apps. For example, in their 2019 paper, Mohr et al [54] found that coaching was associated with more downloads of a mental health app but not long-term engagement with that app. Our results are thus best interpreted as exploratory signals that suggest productive avenues for exploring coaching as well as guidance for understanding the high degree of heterogeneity that must be unpacked in new research efforts. The classification scheme used in this study was created de novo by our team given the state of this literature and can serve as a useful scaffold to create new versions in the future.

Coaching for mental health apps will continue to expand in scope, necessitating an understanding of its therapeutic potential and implementation into care settings. While current efforts around the role remain diverse, they suggest a flexibility necessary to support the evolving digital mental health space and to work across diverse populations and technologies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Code book and coding results.
[XLSX File (Microsoft Excel File), 23 KB - humanfactors_v9i1e28301_app1.xlsx]

References

1. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy: Theory, Research & Practice 1979;16(3):252-260. [doi: 10.1037/h0085885]
2. Ardito RB, Rabellino D. Therapeutic alliance and outcome of psychotherapy: historical excursus, measurements, and prospects for research. Front Psychol 2011 Aug 18;2:270-220 [FREE Full text] [doi: 10.3389/fpsyg.2011.00270] [Medline: 22028698]
3. Carpenter-Song E, Acquilano SC, Noel V, Al-Abdulmunem M, Torous J, Drake RE. Individualized Intervention to Support Mental Health Recovery Through Implementation of Digital Tools into Clinical Care: Feasibility Study. Community Ment Health J 2022 Jan 21;58(1):99-110 [FREE Full text] [doi: 10.1007/s10597-021-00798-6] [Medline: 33611684]
4. Zhai Y. A Call for Addressing Barriers to Telemedicine: Health Disparities during the COVID-19 Pandemic. Psychother Psychosom 2021 Jun 4;90(1):64-66 [FREE Full text] [doi: 10.1159/000509000] [Medline: 32498070]
5. Kim JH, Desai E, Cole M. How The Rapid Shift To Telehealth Leaves Many Community Health Centers Behind During The COVID-19 Pandemic. Health Affairs Forefront. 2020 Feb 06. URL: https://www.healthaffairs.org/do/10.1377/forefront.20200529.449762/full/ [accessed 2021-11-11]
6. Byambasuren O, Beller E, Hoffmann T, Glasziou P. mHealth App Prescription in Australian General Practice: Pre-Post Study. JMIR Mhealth Uhealth 2020 Jun 01;8(6):e16497. [doi: 10.2196/16497] [Medline: 32478660]
7. Mordecai D, Histon T, Neuwirth E, Heisler WS, Kraft A, Bang Y, et al. How Kaiser Permanente Created a Mental Health and Wellness Digital Ecosystem. NEJM Catalyst 2021 Jan 01;2(1):1-20. [doi: 10.1056/CAT.20.0295]
8. Triana AJ, Gusdorf RE, Shah KP, Horst SN. Technology Literacy as a Barrier to Telehealth During COVID-19. Telemed J E Health 2020 Sep 01;26(9):1118-1119. [doi: 10.1089/tmj.2020.0155] [Medline: 32429770]
9. Wisniewski H, Torous J. Digital navigators to implement smartphone and digital tools in care. Acta Psychiatr Scand 2020 Apr 26;141(4):350-355 [FREE Full text] [doi: 10.1111/acps.13149] [Medline: 31930477]
10. Ben-Zeev D, Drake R, Marsch L. Clinical technology specialists. BMJ 2015 Feb 19;350(feb19 15):h945-h945. [doi: 10.1136/bmj.h945] [Medline: 25697164]
11. Noel VA, Carpenter-Song E, Acquilano SC, Torous J, Drake RE. The technology specialist: a 21st century support role in clinical care. NPJ Digit Med 2019;2:61 [FREE Full text] [doi: 10.1038/s41746-019-0137-6] [Medline: 31388565]
12. McEnery C, Lim MH, Knowles A, Rice S, Gleeson J, Howell S, et al. Social anxiety in young people with first-episode psychosis: Pilot study of the EMBRACE moderated online social intervention. Early Interv Psychiatry 2021 Feb 30;15(1):76-86. [doi: 10.1111/eip.12912] [Medline: 31889431]
13. Alvarez-Jimenez M, Rice S, D'Alfonso S, Leicester S, Bendall S, Pryor I, et al. A Novel Multimodal Digital Service (Moderated Online Social Therapy+) for Help-Seeking Young People Experiencing Mental Ill-Health: Pilot Evaluation Within a National Youth E-Mental Health Service. J Med Internet Res 2020 Aug 13;22(8):e17155 [FREE Full text] [doi: 10.2196/17155] [Medline: 32788151]
14. Tønning ML, Kessing LV, Bardram JE, Faurholt-Jepsen M. Methodological Challenges in Randomized Controlled Trials on Smartphone-Based Treatment in Psychiatry: Systematic Review. J Med Internet Res 2019 Oct 27;21(10):e15362 [FREE Full text] [doi: 10.2196/15362] [Medline: 31663859]
15. Tokgöz P, Hrynyschyn R, Hafner J, Schönfeld S, Dockweiler C. Digital Health Interventions in Prevention, Relapse, and Therapy of Mild and Moderate Depression: Scoping Review. JMIR Ment Health 2021 Apr 16;8(4):e26268 [FREE Full text] [doi: 10.2196/26268] [Medline: 33861201]
16. Hur J, Kim B, Park D, Choi S. A Scenario-Based Cognitive Behavioral Therapy Mobile App to Reduce Dysfunctional Beliefs in Individuals with Depression: A Randomized Controlled Trial. Telemed J E Health 2018 Sep;24(9):710-716. [doi: 10.1089/tmj.2017.0214] [Medline: 29323626]
17. Miner A, Kuhn E, Hoffman JE, Owen JE, Ruzek JI, Taylor CB. Feasibility, acceptability, and potential efficacy of the PTSD Coach app: A pilot randomized controlled trial with community trauma survivors. Psychol Trauma 2016 May;8(3):384-392. [doi: 10.1037/tra0000092] [Medline: 27046668]
18. Kuhn E, Kanuri N, Hoffman JE, Garvert DW, Ruzek JI, Taylor CB. A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. J Consult Clin Psychol 2017 Mar;85(3):267-273. [doi: 10.1037/ccp0000163] [Medline: 28221061]
19. Lüdtke T, Pult LK, Schröder J, Moritz S, Bücker L. A randomized controlled trial on a smartphone self-help application (Be Good to Yourself) to reduce depressive symptoms. Psychiatry Res 2018 Nov;269:753-762. [doi: 10.1016/j.psychres.2018.08.113] [Medline: 30273901]
20. Bakker D, Kazantzis N, Rickwood D, Rickard N. A randomized controlled trial of three smartphone apps for enhancing public mental health. Behav Res Ther 2018 Oct;109:75-83. [doi: 10.1016/j.brat.2018.08.003] [Medline: 30125790]
21. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile Sensing and Support for People With Depression: A Pilot Trial in the Wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
22. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013 Feb 07;13(1):1-20. [doi: 10.1186/1471-244x-13-49]

23. Dagöö J, Asplund RP, Bsenko HA, Hjerling S, Holmberg A, Westh S, et al. Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: a randomized controlled trial. J Anxiety Disord 2014 May;28(4):410-417. [doi: 10.1016/j.janxdis.2014.02.003] [Medline: 24731441]
24. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May 01;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
25. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014 Jan 09;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
26. Depp CA, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK, et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015 Mar 15;174:23-30 [FREE Full text] [doi: 10.1016/j.jad.2014.10.053] [Medline: 25479050]
27. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones – the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol. Med 2015 Jul 29;45(13):2691-2704. [doi: 10.1017/s0033291715000410]
28. Ly KH, Topooco N, Cederlund H, Wallin A, Bergström J, Molander O, et al. Smartphone-Supported versus Full Behavioural Activation for Depression: A Randomised Controlled Trial. PLoS One 2015 May 26;10(5):e0126559 [FREE Full text] [doi: 10.1371/journal.pone.0126559] [Medline: 26010890]
29. Moëll B, Kollberg L, Nasri B, Lindefors N, Kaldo V. Living SMART — A randomized controlled trial of a guided online course teaching adults with ADHD or sub-clinical ADHD to use smartphones to structure their everyday life. Internet Interventions 2015 Mar;2(1):24-31. [doi: 10.1016/j.invent.2014.11.004]
30. Ivanova E, Lindner P, Ly KH, Dahlin M, Vernmark K, Andersson G, et al. Guided and unguided Acceptance and Commitment Therapy for social anxiety disorder and/or panic disorder provided via the Internet and a smartphone application: A randomized controlled trial. J Anxiety Disord 2016 Dec;44:27-35. [doi: 10.1016/j.janxdis.2016.09.012] [Medline: 27721123]
31. Hildebrandt T, Michaelides A, Mackinnon D, Greif R, DeBar L, Sysko R. Randomized controlled trial comparing smartphone assisted versus traditional guided self-help for adults with binge eating. Int J Eat Disord 2017 Nov 27;50(11):1313-1322 [FREE Full text] [doi: 10.1002/eat.22781] [Medline: 28960384]
32. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone Cognitive Behavioral Therapy as an Adjunct to Pharmacotherapy for Refractory Depression: Randomized Controlled Trial. J Med Internet Res 2017 Nov 03;19(11):e373 [FREE Full text] [doi: 10.2196/jmir.8602]
33. Ben-Zeev D, Brian RM, Jonathan G, Razzano L, Pashka N, Carpenter-Song E, et al. Mobile Health (mHealth) Versus Clinic-Based Group Intervention for People With Serious Mental Illness: A Randomized Controlled Trial. Psychiatr Serv 2018 Sep 01;69(9):978-985. [doi: 10.1176/appi.ps.201800063] [Medline: 29793397]
34. Boettcher J, Magnusson K, Marklund A, Berglund E, Blomdahl R, Braun U, et al. Adding a smartphone app to internet-based self-help for social anxiety: A randomized controlled trial. Computers in Human Behavior 2018 Oct;87:98-108. [doi: 10.1016/j.chb.2018.04.052]
35. Bucci S, Barrowclough C, Ainsworth J, Machin M, Morris R, Berry K, et al. Actissist: Proof-of-Concept Trial of a Theory-Driven Digital Intervention for Psychosis. Schizophr Bull 2018 Aug 20;44(5):1070-1080 [FREE Full text] [doi: 10.1093/schbul/sby032] [Medline: 29566206]
36. Liang D, Han H, Du J, Zhao M, Hser Y. A pilot study of a smartphone application supporting recovery from drug addiction. J Subst Abuse Treat 2018 May;88:51-58 [FREE Full text] [doi: 10.1016/j.jsat.2018.02.006] [Medline: 29606226]
37. Schlosser D, Campellone TR, Truong B, Etter K, Vergani S, Komaiko K, et al. Efficacy of PRIME, a Mobile App Intervention Designed to Improve Motivation in Young People With Schizophrenia. Schizophr Bull 2018 Aug 20;44(5):1010-1020 [FREE Full text] [doi: 10.1093/schbul/sby078] [Medline: 29939367]
38. Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, et al. A mobile app for social anxiety disorder: A three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. J Consult Clin Psychol 2018 Jun;86(6):493-504. [doi: 10.1037/ccp0000301] [Medline: 29781648]
39. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV. The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial. Psychol. Med 2019 Apr 04;50(5):838-848. [doi: 10.1017/s0033291719000710]
40. Krzystanek M, Borkowski M, Skałacka K, Krysta K. A telemedicine platform to improve clinical parameters in paranoid schizophrenia patients: Results of a one-year randomized study. Schizophr Res 2019 Feb;204:389-396. [doi: 10.1016/j.schres.2018.08.016] [Medline: 30154027]
41. Stiles-Shields C, Montague E, Kwasny MJ, Mohr DC. Behavioral and cognitive intervention strategies delivered via coached apps for depression: Pilot trial. Psychol Serv 2019 May;16(2):233-238 [FREE Full text] [doi: 10.1037/ser0000261] [Medline: 30407055]
42. Teng M, Hou Y, Chang S, Cheng H. Home-delivered attention bias modification training via smartphone to improve attention control in sub-clinical generalized anxiety disorder: A randomized, controlled multi-session experiment. J Affect Disord 2019 Mar 01;246:444-451. [doi: 10.1016/j.jad.2018.12.118] [Medline: 30599367]

43. Enock PM, Hofmann SG, McNally RJ. Attention Bias Modification Training Via Smartphone to Reduce Social Anxiety: A Randomized, Controlled Multi-Session Experiment. Cogn Ther Res 2014 Mar 4;38(2):200-216. [doi: 10.1007/s10608-014-9606-z]
44. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized Controlled Trial of SuperBetter, a Smartphone-Based/Internet-Based Self-Help Tool to Reduce Depressive Symptoms. Games Health J 2015 Jun;4(3):235-246. [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
45. Possemato K, Kuhn E, Johnson E, Hoffman JE, Owen JE, Kanuri N, et al. Using PTSD Coach in primary care with and without clinician support: a pilot randomized controlled trial. Gen Hosp Psychiatry 2016 Jan;38:94-98. [doi: 10.1016/j.genhosppsych.2015.09.005] [Medline: 26589765]
46. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare: An Eclectic, Skills-Based App Suite for the Treatment of Depression and Anxiety. J Med Internet Res 2017 Jan 05;19(1):e10 [FREE Full text] [doi: 10.2196/jmir.6645] [Medline: 28057609]
47. Pfeiffer PN, Henry J, Ganoczy D, Piette JD. Pilot study of psychotherapeutic text messaging for depression. J Telemed Telecare 2016 Jul 21;23(7):665-672. [doi: 10.1177/1357633x16659955]
48. Henson P, Wisniewski H, Hollis C, Keshavan M, Torous J. Digital mental health apps and the therapeutic alliance: initial review. BJPsych Open 2019 Jan;5(1):e15 [FREE Full text] [doi: 10.1192/bjo.2018.86] [Medline: 30762511]
49. Goldberg SB, Baldwin SA, Riordan KM, Torous J, Dahl CJ, Davidson RJ, et al. Alliance With an Unguided Smartphone App: Validation of the Digital Working Alliance Inventory. Assessment 2021 May 18:10731911211015310. [doi: 10.1177/10731911211015310] [Medline: 34000843]
50. Wisniewski H, Gorrindo T, Rauseo-Ricupero N, Hilty D, Torous J. The Role of Digital Navigators in Promoting Clinical Care and Technology Integration into Practice. Digit Biomark 2020 Nov 26;4(Suppl 1):119-135 [FREE Full text] [doi: 10.1159/000510144] [Medline: 33442585]
51. Callejas Z, Griol D, Benghazi K, Noguera M, Chollet G, Torres M, et al. Measuring and Fostering Engagement with Mental Health e-Coaches. Companion Publication of the 2020 International Conference on Multimodal Interaction 2002 Dec 27:275-279. [doi: 10.1145/3395035.3425316]
52. Bradway M, Gabarron E, Johansen M, Zanaboni P, Jardim P, Joakimsen R, et al. Methods and Measures Used to Evaluate Patient-Operated Mobile Health Interventions: Scoping Literature Review. JMIR Mhealth Uhealth 2020 Apr 30;8(4):e16814 [FREE Full text] [doi: 10.2196/16814] [Medline: 32352394]
53. Ng MM, Firth J, Minen M, Torous J. User Engagement in Mental Health Apps: A Review of Measurement, Reporting, and Validity. Psychiatr Serv 2019 Jul 01;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
54. Mohr DC, Schueller SM, Tomasino KN, Kaiser SM, Alam N, Karr C, et al. Comparison of the Effects of Coaching and Receipt of App Recommendations on Depression, Anxiety, and Engagement in the IntelliCare Platform: Factorial Randomized Controlled Trial. J Med Internet Res 2019 Aug 28;21(8):e13609 [FREE Full text] [doi: 10.2196/13609] [Medline: 31464192]

Abbreviations

MOST+: enhanced moderated online social therapy

Edited by A Kushniruk; submitted 28.02.21; peer-reviewed by C Athanasopoulou, D Ben-Zeev, M Paul; comments to author 11.05.21; revised version received 08.09.21; accepted 03.01.22; published 08.03.22.

Please cite as:
Meyer A, Wisniewski H, Torous J
Coaching to Support Mental Health Apps: Exploratory Narrative Review
JMIR Hum Factors 2022;9(1):e28301
URL: https://humanfactors.jmir.org/2022/1/e28301
doi:10.2196/28301
PMID:35258468

Review

Human Factors and Technological Characteristics Influencing the Interaction of Medical Professionals With Artificial Intelligence–Enabled Clinical Decision Support Systems: Literature Review

Michael Knop[1], MSc; Sebastian Weber[1], MSc; Marius Mueller[1], PhD; Bjoern Niehaves[1], PhD

Department of Information Systems, University of Siegen, Siegen, Germany

Corresponding Author:
Michael Knop, MSc
Department of Information Systems
University of Siegen
Kohlbettstrasse 15
Siegen, 57072
Germany
Phone: 49 15755910502
Email: michael.knop@uni-siegen.de

Abstract

Background: The digitization and automation of diagnostics and treatments promise to alter the quality of health care and improve patient outcomes, whereas the undersupply of medical personnel, high workload on medical professionals, and medical case complexity increase. Clinical decision support systems (CDSSs) have been proven to help medical professionals in their everyday work through their ability to process vast amounts of patient information. However, comprehensive adoption is partially disrupted by specific technological and personal characteristics. With the rise of artificial intelligence (AI), CDSSs have become an adaptive technology with human-like capabilities and are able to learn and change their characteristics over time. However, research has not reflected on the characteristics and factors essential for effective collaboration between human actors and AI-enabled CDSSs.

Objective: Our study aims to summarize the factors influencing effective collaboration between medical professionals and AI-enabled CDSSs. These factors are essential for medical professionals, management, and technology designers to reflect on the adoption, implementation, and development of an AI-enabled CDSS.

Methods: We conducted a literature review including 3 different meta-databases, screening over 1000 articles and including 101 articles for full-text assessment. Of the 101 articles, 7 (6.9%) met our inclusion criteria and were analyzed for our synthesis.

Results: We identified the technological characteristics and human factors that appear to have an essential effect on the collaboration of medical professionals and AI-enabled CDSSs in accordance with our research objective, namely, training data quality, performance, explainability, adaptability, medical expertise, technological expertise, personality, cognitive biases, and trust. Comparing our results with those from research on non-AI CDSSs, some characteristics and factors retain their importance, whereas others gain or lose relevance owing to the uniqueness of human-AI interactions. However, only a few (1/7, 14%) studies have mentioned the theoretical foundations and patient outcomes related to AI-enabled CDSSs.

Conclusions: Our study provides a comprehensive overview of the relevant characteristics and factors that influence the interaction and collaboration between medical professionals and AI-enabled CDSSs. Rather limited theoretical foundations currently hinder the possibility of creating adequate concepts and models to explain and predict the interrelations between these characteristics and factors. For an appropriate evaluation of the human-AI collaboration, patient outcomes and the role of patients in the decision-making process should be considered.

(JMIR Hum Factors 2022;9(1):e28639) doi:10.2196/28639

KEYWORDS

artificial intelligence; clinical decision support systems; CDSS; decision-making; diagnostic decision support; human–computer interaction; human–AI collaboration; machine learning; patient outcomes; deep learning; trust; literature review

Introduction

Background

From a global perspective, many health care systems face comprehensive challenges that affect how care is delivered to society. In this regard, several factors increasingly strain care structures, processes, and the actors involved. For instance, demographic changes and the overall aging of society raise age-related health issues and demands [1,2] and introduce further case complexity; for example, in the form of comorbidity [3]. Simultaneously, a shortage of personnel and medical expertise can be discerned in many—often remote and rural—regions, caused by the low attractiveness of jobs in care due to inappropriate compensation and high workload [4], the attractiveness of urban areas and structures [5], the absence of young graduates willing to establish new or continue existing practices [6], or the trend toward centralized care facilities, inter alia [7]. As a result, larger catchment areas develop for providers who have to cope with deficient and inequitably distributed first-hand access to care [8]. Further, on a societal level, detrimental access to care can marginalize lower socioeconomic groups, as a study from the United States suggests [9], impeding the maintenance of comprehensive and inclusive care. Considering the increasing complexity of medical care on the one hand and the decreasing time and personnel resources on the other hand, the need to actively support clinicians at the point of care is growing.

Clinical Decision Support Systems

Representing a promising and widely adopted technology to render processes and decisions more efficient, so-called clinical decision support systems (CDSSs) are software applications capable of catalyzing and informing the process of decision-making of medical professionals [10]. Although applications exist that target the decisional processes of patients, often called decision aids [11] or patient decision support interventions [12], the clinical use of CDSSs remains the primary domain for decision support. Here, the evaluation of performance, adoption, effectiveness, and impact on patient outcomes advances, but still lacks comprehensive approaches [10], including an analysis of relationships among technological characteristics, continual use, and effects on diagnosis and treatment. Nevertheless, the potential of CDSSs to support diagnostic processes leads to their use in other contexts of medicine; for example, primary care [13], and in several different disciplines, from emergency medicine [14] and dermatology [15] to radiology [16]. Aside from diagnostic purposes, CDSSs are used to detect possible inadequate prescriptions of medication [17] or to simulate different treatment strategies and their impact on patient outcomes [18]. Until today, CDSSs have had partial nonadoption for numerous reasons; for example, workflow disturbances or trust deficits, and their adoption is linked to many different factors concerning technology and human-technology interaction [19,20]. In particular, the subjective perception of and attitude toward the CDSS remains a crucial predictor of adoption [21]. This is because the CDSS surpasses the preferably objective description of medical information (eg, in electronic health records) and interprets this information to support clinical interventions [19]. Meanwhile, the comparability of CDSS among different contexts is difficult because of the already-mentioned variation in user groups (patients, physicians, nurses, etc), medical domains (clinical care, primary care, etc), medical disciplines (dermatology, radiology, etc), and purposes (diagnosis, prescription, treatment, etc).

Owing to technological innovations, health care technologies, including CDSSs, are increasingly enabled by artificial intelligence (AI) [22]. The first evaluation of an AI-enabled CDSS promises increased performance and accuracy compared with a conventional CDSS [23]. In addition, clinicians and experts in the field generally expect simplification of organizational processes, such as patient flows, with the advent of AI [24]. Defined as a technology's capability to work in a way that a human perceives as intelligent [25], AI is used on various occasions with regard to CDSS, such as risk prediction for medical complications [26] and adverse drug effects [27]. However, a rigorous and consistent definition of AI is challenging. Therefore, we followed Helm et al [28] and Schuetz and Venkatesh [29] on their emphasis on the adaptive characteristics of AI, meaning that AI-enabled CDSSs are learning entities that change over time while considering their environmental conditions. Consequently, these systems are not deterministic and may provide different outputs from the same input at different times [30]. Compared with medical professionals, AI-enabled systems can outperform human ratings or predictions; for example, concerning the classification of dermal lesions and proliferation [31]. Regarding the adoption of AI-enabled systems in general, ongoing research reports several concerns indicated by clinicians. Although the fear of being replaced appears to depend on the level of knowledge about the concept of AI that the clinicians possess [32], studies report that clinicians fear being biased by the recommendations of AI, resulting in overconfidence and harmful consequences for patients [33]. In addition, clinicians are concerned that AI might increase the threat of data breaches and the associated risks for patients' privacy, as well as legal consequences resulting from treatment errors [34]. Nevertheless, current research suggests an ambivalent perception of AI. Considering the aforementioned concerns and potential hindrances for adoption, clinicians assume that AI-enabled systems might save time and improve the continuous monitoring of patients [35]. Furthermore, research has highlighted that only a few clinicians comprehend the variety of applications of AI and its conceptual nature [34,35]. Differences in the perception of AI; for example, regarding the fear of being replaced [36], emphasize the subjectivity of clinicians' attitude toward AI.

Considering the ambiguity of concerns regarding clinicians' attitudes toward AI, the mentioned hindrances of CDSS adoption and the similarity between concerns associated with AI and CDSS (eg, biased decision-making, legal consequences, or fear of being replaced), an AI-enabled CDSS might actually increase the relevance of perceptive and subjective factors for adoption and their interplay with technological characteristics. During the process of development and evaluation of AI-enabled CDSS, it became apparent that the potential benefits for clinical performance and treatment quality are maximized by human–AI collaboration, rather than by human-AI competition [31].

However, owing to the interactive and adaptive nature of AI-enabled CDSS, traditional theories and models to explain the use and adoption of these systems forfeit their power to explain and predict a successful collaboration between AI and human beings [29,37]. Specific factors regarding AI-enabled technology and human actors such as dermatologists, radiologists, and other medical professionals are emphasized to influence the relationship among them, including the explainability or understandability of the system [38], its purpose [39], and the resulting trust a human actor perceives in the system [40]. Considering that factors related to the subjective attitude and perception of clinicians, such as trust, already impact the adoption of non-AI CDSS [21,41,42], we argue that the advent of AI-enabled systems increases the importance of specific factors that are not exclusively bound to technological characteristics. Considering the already investigated hindrances impeding the adoption of CDSS by clinicians [43,44], the lack of a sound theoretical basis, or the reliance on traditional theoretical approaches within ongoing research [45], the need for a review of AI-specific factors influencing the collaboration between AI and human actors has increased.

Human-AI Interaction and Collaboration

To understand the dyadic relationship between humans and AI, it is necessary to understand key concepts and their interrelations. Although many researchers use the term interaction [46], literature defining what interaction means is seldom. Hornbæk et al [46] showed that there is no common definition and identified 7 concepts of interaction that highlight different perspectives. However, the human–computer interaction framework of Li and Zhang [47] shows that interaction can be generally understood as a process of using a technology for a task in a specific context. In turn, collaboration etymologically stems from the term *collaborare* which means *work with*. As the origin reveals, collaboration can be understood as a joint effort in which a common goal is pursued. From our perspective, collaboration is thus a successful interaction with an adaptive AI-enabled system. Under the assumption that both human and AI-enabled systems are not error-free, a human-AI collaboration is effective when errors are prevented. In this context, a key driver of such effective collaboration is that medical professionals perceive the system as trustworthy (ie, a certain level of trust) for the tasks to be done and accept it. Trust is a complex psychological construct that is described as the will to make oneself vulnerable [48]. If a party considers another party to be trustworthy, the relationship is in turn determined by the perception of the other parties' attributes of ability (the legitimacy of a system's recommendation for a specific decision), benevolence (the accordance of a human actor's and the system's intention and motivation to do good), and integrity (the accordance of a human actor's and the system's superordinate values) [40]. Nevertheless, it remains unclear how system design can influence the perception of trustworthiness and what human traits foster the propensity to trust.

Objectives

The objective of this study is to summarize the factors influencing effective collaboration between medical professionals and AI-enabled CDSSs. Capturing these factors is essential for medical professionals, management, and technology designers to reflect the adoption, implementation, and development of AI-enabled CDSSs [48,49]. Further, we seek to explore what specific outcomes are used to evaluate successful collaboration between humans and AI-enabled CDSSs (performance, effectiveness, impact on patient outcomes, etc) and the theoretical foundations on which they are based. Finally, the comparison between factors that are associated with AI-enabled CDSSs and those associated with CDSSs not enabled by AI appears to be important in evaluating the extent to which the current literature has already reflected the uniqueness of human-AI collaboration.

Methods

Overview

We conducted a narrative review to summarize the current literature regarding our specific objectives [50]. In the following, we report the search for relevant literature to meet our objective, its selection, and its synthesis to counteract the subjectivity of our results [50]. We selected 3 different meta-databases to search for studies that met our research objective. We defined our search strategy in accordance with the relatively broad scope of our study [51]. To report our results, we followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for reviews [52]. Through our initial search, we identified 1161 studies by screening titles and abstracts, of which 100 (8.61%) satisfied our inclusion criteria. Through a backward search, we identified another study that was included in our full-text assessment, resulting in 101 articles assessed for eligibility. Finally, 6.9% (7/101) of studies were included in our synthesis of results.

Databases

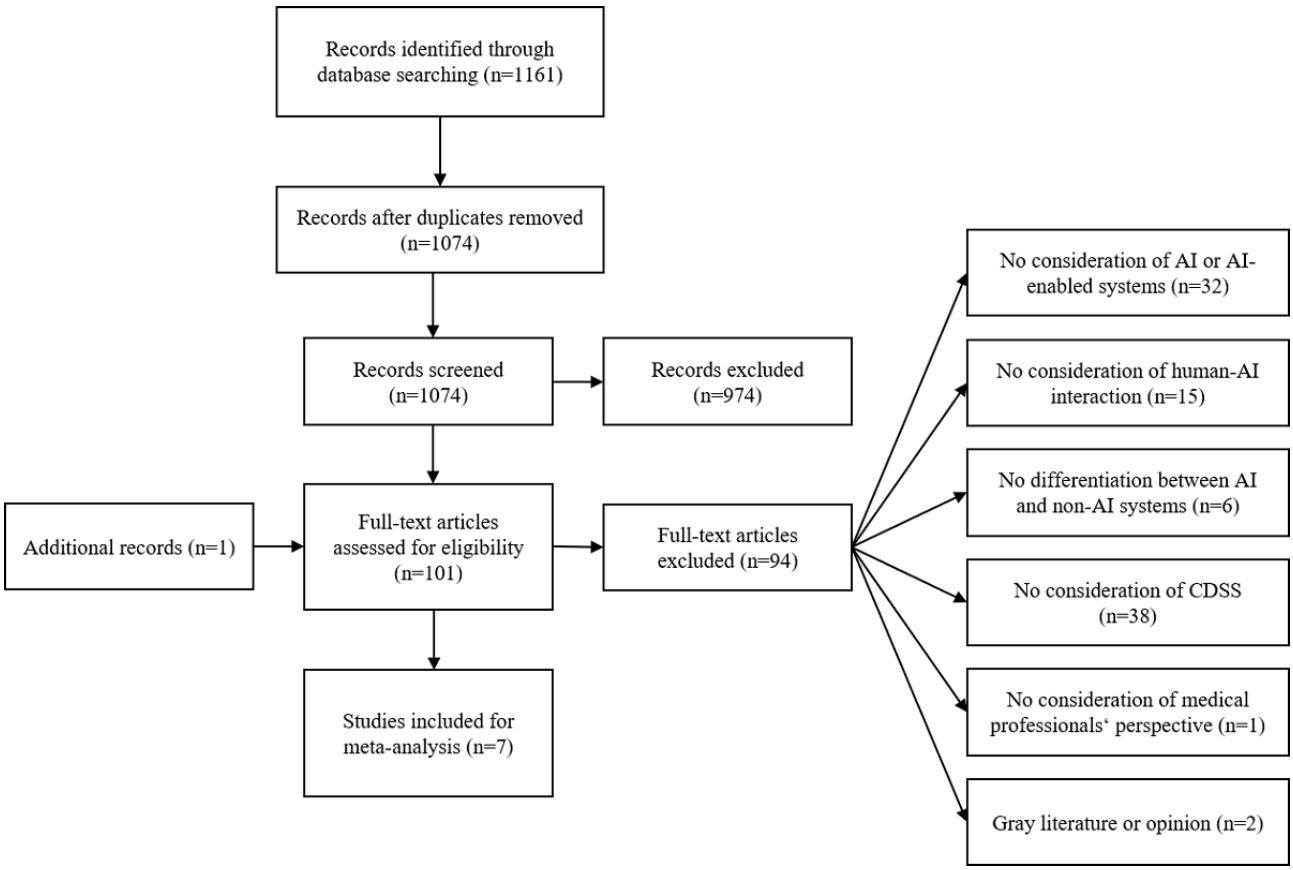
We included the databases PubMed, PsycInfo, and Business Source Complete for our literature review. PubMed indexes >5000 journals in the fields of medicine, health care, and related disciplines. We used PubMed, in particular, to gather information about the clinical effectiveness and implementation of AI-enabled CDSSs. PsycInfo contains >2000 journals from behavioral and social science research. We searched PsycInfo to examine the psychological dimensions of AI-enabled CDSSs and decisional processes. Finally, we scanned results from Business Source Complete, containing >1000 journals in the field of business sciences, to obtain insights regarding our objective from an economic and procedural perspective.

Study Selection

We combined 2 different sections of search terms (AND conditions). The first section represented the technologies associated with the objective of our research *(AI OR artificial intelligence OR machine learning OR cognitive computing OR intelligent agent OR decision support OR recommendation agent)*. The second section reflected the interactional dimension of human-AI collaboration *(trust* OR acceptance OR *agreement OR consent OR compliance OR congruency OR collaboration OR resistance)*. We included articles published in English over the last 10 years. To select relevant literature,

2 authors (MK and SW) independently screened titles and abstracts to exclude articles that did not involve AI-enabled technology (see the definition in the *Clinical Decision Support Systems* section) and those that were not related to health care or medicine. The inclusion and exclusion criteria were discussed in detail before the screening. In addition, to familiarize themselves with the procedure, an initial sample of 100 entries was screened. A high level of agreement was achieved, and disagreements were resolved through discussion between the 2 authors (MK and SW). In the remaining papers, only a few borderline cases were discussed until consensus was reached, and both the authors (MK and SW) finally came to the same result. In the full-text screening, articles that did not involve AI or AI-enabled systems (n=32), did not consider the interaction between the human actors and AI-enabled systems (n=15), did not distinguish between AI-enabled and non–AI-enabled CDSSs (n=6), did not involve CDSS (n=38) or the perspective of medical professionals (n=1), or appeared to be gray literature or opinion (n=2) were excluded. Detailed documentation of the exclusion process for full-text screening is provided in Multimedia Appendix 1, where all excluded studies and the reasons for exclusion are presented. The selection of relevant literature is represented through a PRISMA flowchart (Figure 1). If articles were eligible, we summarized and reported the specific factors influencing effective collaboration.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. AI: artificial intelligence; CDSS: clinical decision support system.



Results

Overview

On the basis of our study selection, 7 studies were included in our final synthesis. From our perspective, this result stems from the fact that many studies of AI-enabled CDSSs (1) compare solely the diagnostic accuracies of human raters and those of AI-enabled systems and (2) focus on technological characteristics and the development of these systems, but do not discuss their effects on the interaction or collaboration between technology and human actors. Therefore, most (5/7, 71%) of the included studies reflected on the relevance of specific characteristics or factors by contemplating the objective from a meta-perspective (Table 1).

Table 1. Summary of study characteristics included in our review.

Study	Type of study	Context	Focal point of interest
Cabitza et al [53]	Narrative review	Clinical care; health care (general); clinicians; no specific purpose	Trust
Felmingham et al [54]	Narrative review	Clinical care; dermatology; physicians; diagnostics	Mortality or morbidity
Gomolin at al [55]	Narrative review	Clinical care; dermatology; physicians; diagnostics	Explainability
Reyes et al [56]	Narrative review	Clinical care; radiology; physicians; diagnostics	Trust; explainability
Jeng and Tzeng [57]	Quantitative study	Clinical care; health care (general); physicians; diagnostics	Intention
Tschandl et al [31]	Quantitative study	Clinical care; dermatology; physicians; diagnostics	Performance
Asan et al [30]	Narrative review	Clinical care; health care (general); clinicians; no specific purpose	Trust

Factors Influencing Collaboration

Technological Characteristics

This study addresses different dimensions or steps in the development, implementation, and adoption of an AI-enabled CDSS. The technological characteristics of these systems, that is, the abilities and attributes of technology that are defined by their design [58], are described as meaningful determinants for the way the interaction between the system and the human actor is shaped. For instance, Cabitza et al [53] concluded that a "truthful, reliable, and representative" system needs high-quality data based on which it is trained. Similarly, Asan et al [30] argued that the development of a "healthy trust relationship" with algorithmic decision-making relies on the thoughtful design of system characteristics. In general, the resulting performance of the system and its ability to explain or justify its conclusions appear to be strong predictors of a positive relationship [30,31,54-56]. Reyes et al [56] defined the explainability of an AI-enabled system as the ability to ensure that a human actor understands "the link between the features used by the machine learning system and the prediction." In current literature, explainability and transparency of a system are often used interchangeably [54] or in the sense that transparency appears to be a superordinate category of explainability [30]. Closely linked to a system's ability to explain its internal processes is the resulting effect on human actors with respect to the subjective interpretability of the given information [55].

Furthermore, Tschandl et al [31] argue that the output of an AI-enabled CDSS in its dimensions of simplicity, granularity, and concreteness might affect the final decision of clinicians; the better an AI-enabled system's output is adapted to the situational context of its use, the more precise the overall diagnostic performance of the AI and humans (eg, clinicians facing multiclass diagnostic problems are supported by AI-based multiclass probabilities). In addition, a study mentioned the importance of usability and user satisfaction for effective human-AI collaboration [53] but does not provide a definition in the context of AI-enabled CDSSs.

Human Factors

In addition, social (eg, trust), psychological (eg, personality traits), and cognitive characteristics (eg, cognitive biases) of a human actor affecting their interaction with technology, that is, human factors [59], appear to be meaningful prerequisites for the relationship between systems and actors as well. Asan et al [30], Tschandl et al [31], Felmingham et al [54], and Jeng and Tzeng [57] argued that the clinical experience of medical professionals is a highly important factor in determining the interaction and performance of human-AI collaboration. In general, these studies show that less experienced physicians benefit the most from AI-enabled CDSSs and attain a higher overall diagnostic accuracy, whereas an experienced physician's diagnostic accuracy differs little or not at all. In addition, Asan et al [30] and Felmingham et al [54] argue that technological experience and even the personality of medical professionals are important factors for medical professionals' decision-making processes, although no study has yet investigated their effect on the collaboration between AI-enabled CDSSs and human actors. Furthermore, Asan et al [30] and Felmingham et al [54] mentioned that the relationship between the system and human actor can be disrupted by several cognitive biases affecting collaboration at different times, that is, confirmation bias, anchoring effect, overconfidence, availability bias, framing effect, premature closure, and automation bias. Already known from medical decision-making in general, cognitive biases alter rational processes of medical professionals, resulting in erroneous diagnostics and treatments [60]. Because of biased thinking in decisional processes and the variety of biases occurring at different times in these processes, AI-enabled CDSSs are prone to be affected by these biases [54].

Among the included studies, a human actor's trust in an AI-enabled CDSS appeared to be another important factor that directly influenced the quality of collaboration and adoption of technology. For instance, Cabitza et al [53] argued that a lack of trust might result from different technological characteristics and their situational fit but always negatively impacts the overall performance of the human-AI team. Reyes et al [56] hypothesized that the comprehensible explainability of a system ensures a high level of trust, including a system's ability to explicate its learning process and essential or most effective determinants for its prediction, as well as adequate and situational visualization of its internal processes. Felmingham et al [54] argued that trust is created through an interactional process between AI and humans. Accordingly, Asan et al [30] also highlighted the interdependency of human factors and system features as constituting factors of trust. However, Asan et al [30] argued that maximizing trust should not be the ultimate goal, as AI also has its limitations in that blind trust could lead

to undesirable consequences. Instead, system designers should establish mechanisms that encourage reciprocal skepticism, create healthy trust relationships, and maximize the accuracy of clinical decisions. From this perspective, trust is highly dependent on the personality of the human actor, system design, and cognitive biases that might emerge in the collaboration. The reported technological characteristics and human factors influencing effective AI-human collaboration are summarized in Table 2.

Table 2. Technological characteristics and human factors influencing and shaping the relationship and collaboration between AI-enabled clinical decision support systems (CDSSs) and human actors.

Parameters	Definition	Study
Technological characteristics		
Training data quality	Information used for training of AI-enabled CDSSs to create a truthful, reliable, and representative system	[53]
Performance	The accuracy and reliability of an AI-enabled CDSS	[30,55]
Explainability or transparency	An AI-enabled CDSS' ability to ensure that a human actor understands the processes that lead to the prediction and the prediction itself	[30,31,54-56]
Adapted output or adaptability	The degree to which an AI-enabled CDSS fits into a specific context or environment according to the subdimensions simplicity, granularity, and concreteness	[31]
Human factors		
Medical expertise	The degree of medical experience of a human actor within the context of collaboration with an AI-enabled CDSS	[30,31,54,57]
Technological expertise	The degree of technological experience of a human actor with regard to an AI-enabled CDSS	[30,54]
Personality	A medical professional's attributes and characteristics that influence the interaction with AI-enabled a CDSS	[54]
Cognitive biases	The cognitive processes that alter rational decision-making and perceptions of an AI-enabled CDSS	[30,54]
Trust	The subjective impression of a medical professional that an AI-enabled CDSS is truthful and reliable	[30,53,54]

Evaluation of Medical Outcomes

Of the 7 included studies, only 1 (14%) study mentioned the interrelation between an effective human-AI collaboration and primary clinical outcomes. Reviewing an AI-enabled CDSS for skin cancer diagnostics, Felmingham et al [54] mentioned the possible impacts of these systems on a patient's morbidity and mortality associated with skin cancer in general. Other studies described secondary outcomes, such as a system's mathematical accuracy [55] or behavioral intentions to use a CDSS [57]. No study investigated the impact of technological characteristics or human factors on primary clinical outcomes.

Theoretical Foundation of Research

Of the 7 included studies, only 1 (14%) study mentioned the theoretical foundations on which implications for practice are based explicitly. Jeng and Tzeng [57] derived hypotheses for their empirical investigation from the unified theory of acceptance and use of technology, which is a technology acceptance theory widely adopted to explain the intention to use technology and the subsequent use behavior [61]. An important predecessor in this theoretical model is social influence (ie, "...the degree to which an individual perceives that important others believe he or she should use the new system" [61]). However, based on their results, Jeng and Tzeng [57] discarded their theoretical assumption about social influence affecting clinicians' intentions to use a CDSS. Felmingham et al [54] discussed the role of cognitive biases in decisional processes involving AI-enabled CDSSs. Nevertheless, Felmingham et al [54] did not explicitly mention the origin of cognitive biases in the prospect theory by Kahneman and Tversky [62].

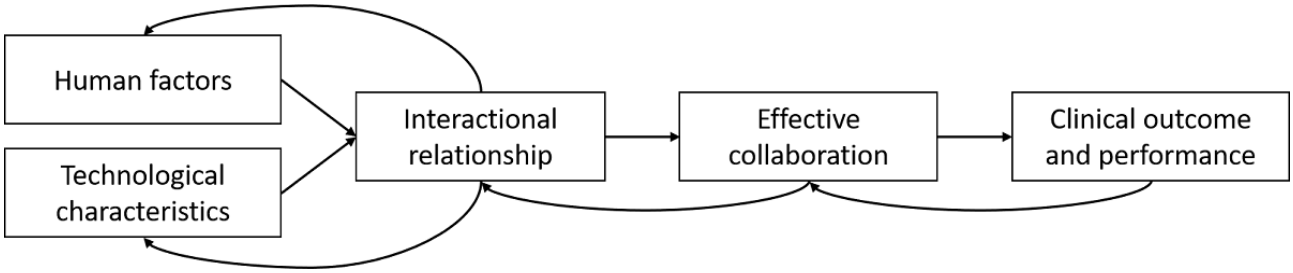
Discussion

Principal Findings

Our results show that only a few (7/101, 6.9%) studies have already broached the issue of individual factors influencing effective collaboration between a human actor and an AI-enabled CDSS. Although unique considerations with regard to these systems appear; for example, the important role of trust [30,53], scarce empirical evidence exists for the relational structure of essential factors or characteristics. In addition, many studies did not describe the involved system and its characteristics extensively, enabling differentiation between AI and non–AI-enabled systems accurately [42]. Therefore, we argue that a more thorough description of the involved system and its characteristics is highly relevant for future research as it lays the foundation for comparing different systems and their effectiveness. Nevertheless, in the process of reviewing the literature, we were able to differentiate between factors primarily associated with technological structures and functions (technological characteristics), and those primarily associated with human actors' psychological or perceptual attributes (human factors). Both technological characteristics and human factors influence the nature of the interaction between human actors and AI-enabled CDSSs. Interestingly, some technological

characteristics and human factors appear to be antecedents of interaction; for example, the personality of medical professionals [54], whereas others appear to be effects of an interaction [53]. Therefore, as suggested by Felmingham et al [54], it can be assumed that human factors and technological characteristics are mutually dependent and together shape the interaction between human actors and AI-enabled CDSSs. As described in the *Background* section, the shape of an interaction between human actors and AI and their resulting interactional relationship can be considered a condition for successful collaboration. However, the foundation for evaluating an AI-enabled CDSS differs, in accordance with current research addressing non-AI CDSS [20]. Studies from our results discussed the accuracy or mathematical performance of systems, adoption by medical professionals, sustainability and congruency of interaction, and the effects on patient outcomes to be relevant for evaluation. Although the effectiveness of a collaboration between human actors and AI currently depends on the context and objective of a system [53], the paradigm of medicine clearly dictates the final evaluation of a CDSS by its ability to improve primary and secondary outcomes of patients [22]. As AI-enabled systems are characterized by their adaptive nature [29], processes of individual interaction and collaboration are likely to be iterative and reciprocal and will change and be refined over time. Figure 2 summarizes this process based on our results and can be considered a proposed descriptive framework for human-AI collaboration.

Figure 2. Steps and elements of reciprocal processes of human–artificial intelligence collaboration.



When comparing our results to research concerning medical professionals' interactions with non-AI CDSS, high correspondence can be noted. Khairat et al [20] mentioned workflow fit (adaptability), computer literacy (technological expertise), trust, a general optimistic attitude of clinicians (personality), and clinical expertise (medical expertise) as important factors for effective adoption. In addition, Khairat et al [20] reported usability and perceived usefulness as determinants. As perceived usefulness needs further concretization within the context of AI-enabled systems [53], usability might generate only minor relevance for AI-enabled CDSSs, as these systems are based on automated processes of use and integrate human-like ways to communicate (eg, natural language processing for voice control) [29]. In contrast, the explainability of a CDSS appears to be a technological characteristic that strongly influences the collaboration between humans and the system, whether it is enabled by AI or not [63]. However, differentiation of explainability and related terms such as understandability, interpretability, and transparency has not yet been completed, and the impact of explainability on other relevant factors, including trust, has not yet been empirically verified [63]. In general, it is not clear how and if technological characteristics and human factors influence other specific aspects of collaboration between human actors and AI-enabled CDSSs. For instance, studies suggest that high clinical expertise influences overall collaborative performance [54] but does not hypothesize possible explanations. Clinical expertise might be associated with a lack of trust in these systems, overconfidence biases, or the fact that these systems are sometimes less accurate than experienced physicians.

Furthermore, other studies involving non-AI CDSSs have emphasized the essential role of trust in the effective interaction between humans and the system. Trust is a multidimensional construct. A lack of trust might result from reservations regarding the mathematical accuracy or appropriateness of a system or the purpose of a system in improving patient outcomes [64]. As our literature review reveals the importance of the technological accuracy of AI-enabled CDSSs, research focusing on trust in human-like technology has shown that ability, benevolence, and integrity are essential prerequisites for sustainable adoption [37,40]. However, only 14% (1/7) of the included studies highlighting trust in its role in successful collaboration defined the actual meaning of trust [30], and none of the included studies paid attention to the prerequisites. Considering the inconsistent definition of trust in technology [48,65], future research might reveal important prerequisites for trust within the interaction between human actors and AI-enabled technologies. In addition, the relationship between trust in AI-enabled CDSSs and improvement in clinical outcomes requires further investigation.

Findings from our literature as well as ongoing research concerning non-AI patient decisional aid suggest that a stronger theoretical foundation for the interaction between human actors and CDSSs is important [66]. Felmingham et al [54] already demonstrated that cognitive biases, originating from the prospect theory, might decisively impact effective collaboration, that is, the tendency to confirm assumptions already made rather than falsify them, known as confirmation bias [67], might distort the relationship between medical professionals and AI-enabled CDSS in the sense that they might not accept a different opinion except their own. In contrast, relying on automated information instead of vigilantly seeking and interpreting information, known as automation bias [68], might actually cause the unreflected acceptance of suggestions made by a CDSS. Therefore, to discuss suitable theoretical foundations, it might be helpful to further explicate and structure the aforementioned nontransparent relations of different constructs, factors, and characteristics influencing decision-making and collaboration.

In addition, problems originating from the application of traditional technology-centered theories (such as the technology acceptance model or unified theory of acceptance and use of technology) on AI-enabled decision-making might lead to inappropriate results [29,69]. Theories concerning the trust-based adoption of human-like technology [40] promise to encounter these deficits by emphasizing the interactional components of technology adoption and use.

Limitations

Our study had some limitations. As some studies derived their conclusions about collaboration between AI-enabled CDSSs and human actors from studies of CDSSs not enabled by AI or assigned results from non-CDSSs to CDSSs, reasoning about interrelations between different technological characteristics and human factors is preliminary and requires further investigation. Although our results fit well with the current findings about the uniqueness and specific nature of human-AI interaction, very few (7/101, 6.9%) studies, of which most were narrative reviews, were included because of our innovative and novel objective as well as the specific context. This may be a result of our relatively narrow search, which could be extended by explicating the related constructs and prerequisites of trust. Explorative empirical studies based on suitable theoretical foundations might yield frameworks and models to structure future research on AI-enabled CDSSs, as our study primarily provides an orientation about relevant individual characteristics and factors. The consideration of environmental influences (eg, organizational policies or culture [30] and patients' views [70]) on AI-supported decisional processes for medical care is vital for a comprehensible understanding but cannot be provided within the scope of our review.

Conclusions

We extracted the technological characteristics and human factors relevant for effective collaboration between medical professionals and AI-enabled CDSSs. Although most of the findings from previous research on non–AI-enabled CDSSs are in accordance with our results, the weighting of specific factors might change with AI-enabled systems. The adaptive and increasing human-like nature of AI-enabled CDSSs emphasizes the time sensitivity and reciprocity of decisional processes that should ultimately lead to an improvement in care. Cognitive biases may occur at any time during these processes, varying the effectiveness of collaboration. Explainability remains an essential prerequisite for interaction, and the expertise and personalities of medical professionals have come into focus. In addition, trust between humans and the system emerges as a central aspect of decisional support, whereas the interrelations among these facets still need to be investigated. Concepts such as shared decision-making justify the integration of patients' demands and wishes, an important factor for medical care, and its role in human-AI collaboration is yet underrepresented. Currently, it is unclear how these concepts can be integrated into AI-enhanced decisional processes and to what extent medical decisions with the help of the CDSS are influenced by the subjective meaning and understanding of diagnoses or treatments by patients. In addition, as several studies have measured the effectiveness of collaboration by means of other parameters, primary and secondary patient outcomes should be considered in future research.

As described earlier, modern health care structures are under increasing pressure. Involved medical professionals face immense workloads per capita, and the supply of personnel declines. Because these structures form the initial access points for most citizens in need of care and treatment, approaches that foster more efficient decision-making and treatment processes are becoming imperative to maintain comprehensive care. Thus, an AI-enabled CDSS represents an important and future-oriented measure that enables actors in the health care domain to improve resource allocation, make timelier and less stressful decisions, and cope with shortages in personnel, facilities, and expertise. However, the potential application of CDSSs and pursued benefits calls for investigations that shed light on how AI-enabled processes can be implemented within prevalent health care structures so that the associated risks and challenges, such as the oversimplification of individual patient data or the automated initiation of suboptimal or erroneous treatments, can be mitigated.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Documentation of full-text exclusion.
[XLSX File (Microsoft Excel File), 25 KB - humanfactors_v9i1e28639_app1.xlsx]

References

1. Davis S, Bartlett H. Healthy ageing in rural Australia: issues and challenges. Australas J Ageing 2008;27(2):56-60. [doi: 10.1111/j.1741-6612.2008.00296.x] [Medline: 18713193]
2. Demiris G, Hensel BK. Technologies for an aging society: a systematic review of "smart home" applications. Yearb Med Inform 2008:33-40. [Medline: 18660873]
3. Kazdin AE, Whitley MK. Comorbidity, case complexity, and effects of evidence-based treatment for children referred for disruptive behavior. J Consult Clin Psychol 2006;74(3):455-467. [doi: 10.1037/0022-006X.74.3.455] [Medline: 16822103]

4. Thommasen HV, Lavanchy M, Connelly I, Berkowitz J, Grzybowski S. Mental health, job satisfaction, and intention to relocate. Opinions of physicians in rural British Columbia. Can Fam Physician 2001;47:737-744 [FREE Full text] [Medline: 11340754]
5. Yang J. Potential urban-to-rural physician migration: the limited role of financial incentives. Can J Rural Med 2003;8(2):101-106.
6. Adarkwah CC, Schwaffertz A, Labenz J, Becker A, Hirsch O. Assessment of the occupational perspectives of general practitioners in a rural area. Results from the study HaMedSi (Hausärzte [GPs] for Medical education in Siegen-Wittgenstein). MMW Fortschr Med 2019;161(Suppl 6):9-14. [doi: 10.1007/s15006-019-0919-4] [Medline: 31587169]
7. Rechel B, Džakula A, Duran A, Fattore G, Edwards N, Grignon M, et al. Hospitals in rural or remote areas: an exploratory review of policies in 8 high-income countries. Health Policy 2016;120(7):758-769 [FREE Full text] [doi: 10.1016/j.healthpol.2016.05.011] [Medline: 27312144]
8. Politzer RM, Yoon J, Shi L, Hughes RG, Regan J, Gaston MH. Inequality in America: the contribution of health centers in reducing and eliminating disparities in access to care. Med Care Res Rev 2001;58(2):234-248. [doi: 10.1177/107755870105800205] [Medline: 11398647]
9. Baah FO, Teitelman AM, Riegel B. Marginalization: conceptualizing patient vulnerabilities in the framework of social determinants of health-an integrative review. Nurs Inq 2019;26(1):e12268 [FREE Full text] [doi: 10.1111/nin.12268] [Medline: 30488635]
10. Kruse CS, Ehrbar N. Effects of computerized decision support systems on practitioner performance and patient outcomes: systematic review. JMIR Med Inform 2020;8(8):e17283 [FREE Full text] [doi: 10.2196/17283] [Medline: 32780714]
11. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017;4:CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
12. Elwyn G, Scholl I, Tietbohl C, Mann M, Edwards AG, Clay C, et al. "Many miles to go …": a systematic review of the implementation of patient decision support interventions into routine clinical practice. BMC Med Inform Decis Mak 2013;13 Suppl 2:S14 [FREE Full text] [doi: 10.1186/1472-6947-13-S2-S14] [Medline: 24625083]
13. Nurek M, Kostopoulou O, Delaney BC, Esmail A. Reducing diagnostic errors in primary care. A systematic meta-review of computerized diagnostic decision support systems by the LINNEAUS collaboration on patient safety in primary care. Eur J Gen Pract 2015;21 Suppl(sup1):8-13 [FREE Full text] [doi: 10.3109/13814788.2015.1043123] [Medline: 26339829]
14. Bennett P, Hardiker NR. The use of computerized clinical decision support systems in emergency care: a substantive review of the literature. J Am Med Inform Assoc 2017;24(3):655-668 [FREE Full text] [doi: 10.1093/jamia/ocw151] [Medline: 28031285]
15. Anderson AM, Matsumoto M, Saul MI, Secrest AM, Ferris LK. Accuracy of skin cancer diagnosis by physician assistants compared with dermatologists in a large health care system. JAMA Dermatol 2018;154(5):569-573 [FREE Full text] [doi: 10.1001/jamadermatol.2018.0212] [Medline: 29710082]
16. Lambin P, Leijenaar RT, Deist TM, Peerlings J, de Jong EE, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14(12):749-762. [doi: 10.1038/nrclinonc.2017.141] [Medline: 28975929]
17. Scott IA, Pillans PI, Barras M, Morris C. Using EMR-enabled computerized decision support systems to reduce prescribing of potentially inappropriate medications: a narrative review. Ther Adv Drug Saf 2018;9(9):559-573 [FREE Full text] [doi: 10.1177/2042098618784809] [Medline: 30181862]
18. Valdes G, Simone 2nd CB, Chen J, Lin A, Yom SS, Pattison AJ, et al. Clinical decision support of radiotherapy treatment planning: a data-driven machine learning strategy for patient-specific dosimetric decision making. Radiother Oncol 2017;125(3):392-397. [doi: 10.1016/j.radonc.2017.10.014] [Medline: 29162279]
19. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17 [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
20. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
21. Laka M, Milazzo A, Merlin T. Factors that impact the adoption of clinical decision support systems (CDSS) for antibiotic management. Int J Environ Res Public Health 2021;18(4):1901 [FREE Full text] [doi: 10.3390/ijerph18041901] [Medline: 33669353]
22. Morley J, Machado CC, Burr C, Cowls J, Joshi I, Taddeo M, et al. The ethics of AI in health care: a mapping review. Soc Sci Med 2020;260:113172. [doi: 10.1016/j.socscimed.2020.113172] [Medline: 32702587]
23. Hameed BM, Shah M, Naik N, Singh Khanuja H, Paul R, Somani BK. Application of artificial intelligence-based classifiers to predict the outcome measures and stone-free status following percutaneous nephrolithotomy for staghorn calculi: cross-validation of data and estimation of accuracy. J Endourol 2021;35(9):1307-1313. [doi: 10.1089/end.2020.1136] [Medline: 33691473]

24. Dawoodbhoy FM, Delaney J, Cecula P, Yu J, Peacock I, Tan J, et al. AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. Heliyon 2021;7(5):e06993 [FREE Full text] [doi: 10.1016/j.heliyon.2021.e06993] [Medline: 34036191]
25. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth summer research project on artificial intelligence. Stanford University. 1995. URL: http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf [accessed 2020-12-08]
26. Stone EL. Clinical decision support systems in the emergency department: opportunities to improve triage accuracy. J Emerg Nurs 2019;45(2):220-222. [doi: 10.1016/j.jen.2018.12.016] [Medline: 30846145]
27. Patterson BW, Pulia MS, Ravi S, Hoonakker PL, Schoofs Hundt A, Wiegmann D, et al. Scope and influence of electronic health record-integrated clinical decision support in the emergency department: a systematic review. Ann Emerg Med 2019;74(2):285-296 [FREE Full text] [doi: 10.1016/j.annemergmed.2018.10.034] [Medline: 30611639]
28. Helm JM, Swiergosz AM, Haeberle HS, Karnuta JM, Schaffer JL, Krebs VE, et al. Machine learning and artificial intelligence: definitions, applications, and future directions. Curr Rev Musculoskelet Med 2020;13(1):69-76 [FREE Full text] [doi: 10.1007/s12178-020-09600-8] [Medline: 31983042]
29. Schuetz S, Venkatesh V. Research perspectives: the rise of human machines: how cognitive computing systems challenge assumptions of user-system interaction. J Assoc Inf Syst 2020;21(2):460-482. [doi: 10.17705/1jais.00608]
30. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res 2020;22(6):e15154 [FREE Full text] [doi: 10.2196/15154] [Medline: 32558657]
31. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human-computer collaboration for skin cancer recognition. Nat Med 2020;26(8):1229-1234. [doi: 10.1038/s41591-020-0942-0] [Medline: 32572267]
32. Abdullah R, Fakieh B. Health care employees' perceptions of the use of artificial intelligence applications: survey study. J Med Internet Res 2020;22(5):e17620 [FREE Full text] [doi: 10.2196/17620] [Medline: 32406857]
33. Liyanage H, Liaw ST, Jonnagaddala J, Schreiber R, Kuziemsky C, Terry AL, et al. Artificial intelligence in primary health care: perceptions, issues, and challenges. Yearb Med Inform 2019;28(1):41-46 [FREE Full text] [doi: 10.1055/s-0039-1677901] [Medline: 31022751]
34. Castagno S, Khalifa M. Perceptions of artificial intelligence among healthcare staff: a qualitative survey study. Front Artif Intell 2020;3:578983 [FREE Full text] [doi: 10.3389/frai.2020.578983] [Medline: 33733219]
35. Laï MC, Brian M, Mamzer MF. Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France. J Transl Med 2020;18(1):14 [FREE Full text] [doi: 10.1186/s12967-019-02204-y] [Medline: 31918710]
36. Oh S, Kim JH, Choi SW, Lee HJ, Hong J, Kwon SH. Physician confidence in artificial intelligence: an online mobile survey. J Med Internet Res 2019;21(3):e12422 [FREE Full text] [doi: 10.2196/12422] [Medline: 30907742]
37. Siau K, Wang W. Building trust in artificial intelligence, machine learning, and robotics. Cut Bus Technol J 2018;31(2):47-53.
38. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov 2019;9(4):e1312 [FREE Full text] [doi: 10.1002/widm.1312] [Medline: 32089788]
39. Toreini E, Aitken M, Coopamootoo K, Elliott K, Gonzalez Zelaya C, van Moorsel A. The relationship between trust in AI and trustworthy machine learning technologies. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. New York, NY: Association for Computing Machinery; 2020 Presented at: FAT* '20; January 27-30, 2020; Barcelona, Spain p. 272-283. [doi: 10.1145/3351095.3372834]
40. Lankton NK, McKnight DH, Tripp J. Technology, humanness, and trust: rethinking trust in technology. J Assoc Inf Syst 2015;16(10):880-918. [doi: 10.17705/1jais.00411]
41. Kortteisto T, Komulainen J, Mäkelä M, Kunnamo I, Kaila M. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. BMC Health Serv Res 2012;12:349 [FREE Full text] [doi: 10.1186/1472-6963-12-349] [Medline: 23039113]
42. Johnson MP, Zheng K, Padman R. Modeling the longitudinality of user acceptance of technology with an evidence-adaptive clinical decision support system. Decis Support Syst 2014;57:444-453. [doi: 10.1016/j.dss.2012.10.049]
43. Liberati EG, Ruggiero F, Galuppo L, Gorli M, González-Lorenzo M, Maraldi M, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. Implement Sci 2017;12(1):113 [FREE Full text] [doi: 10.1186/s13012-017-0644-2] [Medline: 28915822]
44. Moxey A, Robertson J, Newby D, Hains I, Williamson M, Pearson SA. Computerized clinical decision support for prescribing: provision does not guarantee uptake. J Am Med Inform Assoc 2010;17(1):25-33 [FREE Full text] [doi: 10.1197/jamia.M3170] [Medline: 20064798]
45. Khong PC, Holroyd E, Wang W. A critical review of the theoretical frameworks and the conceptual factors in the adoption of clinical decision support systems. Comput Inform Nurs 2015;33(12):555-570. [doi: 10.1097/CIN.0000000000000196] [Medline: 26535769]
46. Hornbæk K, Mottelson A, Knibbe J, Vogel D. What do we mean by “Interaction”? An analysis of 35 years of CHI. ACM Trans Comput-Hum Interact 2019;26(4):1-30. [doi: 10.1145/3325285]
47. Li NL, Zhang P. The intellectual development of human-computer interaction research: a critical assessment of the MIS literature (1990-2002). J Assoc Inf Syst 2005;6(11):227-292. [doi: 10.17705/1jais.00070]

48. Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. Acad Manag Rev 1995;20(3):709-734. [doi: 10.2307/258792]
49. Mcknight DH, Carter M, Thatcher JB, Clay PF. Trust in a specific technology: an investigation of its components and measures. ACM Trans Manage Inf Syst 2011;2(2):1-25. [doi: 10.1145/1985347.1985353]
50. Rocco TS, Plakhotnik MS. Literature reviews, conceptual frameworks, and theoretical frameworks: terms, functions, and distinctions. Hum Resour Dev Rev 2009;8(1):120-130. [doi: 10.1177/1534484309332617]
51. Aida A, Svensson T, Svensson AK, Chung UI, Yamauchi T. eHealth delivery of educational content using selected visual methods to improve health literacy on lifestyle-related diseases: literature review. JMIR Mhealth Uhealth 2020;8(12):e18316 [FREE Full text] [doi: 10.2196/18316] [Medline: 33295296]
52. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009;62(10):1006-1012. [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
53. Cabitza F, Campagner A, Balsano C. Bridging the "last mile" gap between AI implementation and operation: "data awareness" that matters. Ann Transl Med 2020;8(7):501 [FREE Full text] [doi: 10.21037/atm.2020.03.63] [Medline: 32395545]
54. Felmingham CM, Adler NR, Ge Z, Morton RL, Janda M, Mar VJ. The importance of incorporating human factors in the design and implementation of artificial intelligence for skin cancer diagnosis in the real world. Am J Clin Dermatol 2021;22(2):233-242. [doi: 10.1007/s40257-020-00574-4] [Medline: 33354741]
55. Gomolin A, Netchiporouk E, Gniadecki R, Litvinov IV. Artificial intelligence applications in dermatology: where do we stand? Front Med (Lausanne) 2020;7:100 [FREE Full text] [doi: 10.3389/fmed.2020.00100] [Medline: 32296706]
56. Reyes M, Meier R, Pereira S, Silva CA, Dahlweid FM, von Tengg-Kobligk H, et al. On the interpretability of artificial intelligence in radiology: challenges and opportunities. Radiol Artif Intell 2020;2(3):e190043 [FREE Full text] [doi: 10.1148/ryai.2020190043] [Medline: 32510054]
57. Jeng DJ, Tzeng GH. Social influence on the use of clinical decision support systems: revisiting the unified theory of acceptance and use of technology by the fuzzy DEMATEL technique. Comput Ind Eng 2012;62(3):819-828. [doi: 10.1016/j.cie.2011.12.016]
58. Cheng YM. Towards an understanding of the factors affecting m-learning acceptance: roles of technological characteristics and compatibility. Asia Pac Manag Rev 2015;20(3):109-119. [doi: 10.1016/j.apmrv.2014.12.011]
59. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
60. Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. Acad Med 2017;92(1):23-30. [doi: 10.1097/ACM.0000000000001421] [Medline: 27782919]
61. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
62. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979;47(2):263-292. [doi: 10.2307/1914185]
63. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. J Biomed Inform 2021;113:103655 [FREE Full text] [doi: 10.1016/j.jbi.2020.103655] [Medline: 33309898]
64. Sousa VE, Lopez KD, Febretti A, Stifter J, Yao Y, Johnson A, et al. Use of simulation to study nurses' acceptance and nonacceptance of clinical decision support suggestions. Comput Inform Nurs 2015;33(10):465-472 [FREE Full text] [doi: 10.1097/CIN.0000000000000185] [Medline: 26361268]
65. Mcknight DH, Chervany NL. The meanings of trust. Minneapolis, MN: Carlson School of Management, University of Minnesota; 1996.
66. Elwyn G, Stiel M, Durand MA, Boivin J. The design of patient decision support interventions: addressing the theory-practice gap. J Eval Clin Pract 2011;17(4):565-574. [doi: 10.1111/j.1365-2753.2010.01517.x] [Medline: 20698918]
67. Nickerson RS. Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol 1998;2(2):175-220. [doi: 10.1037/1089-2680.2.2.175]
68. Mosier KL, Skitka LJ, Heers S, Burdick M. Automation bias: decision making and performance in high-tech cockpits. Int J Aviat Psychol 1997;8(1):47-63. [doi: 10.1207/s15327108ijap0801_3] [Medline: 11540946]
69. Jun S, Plint AC, Campbell SM, Curtis S, Sabir K, Newton AS. Point-of-care cognitive support technology in emergency departments: a scoping review of technology acceptance by clinicians. Acad Emerg Med 2018;25(5):494-507 [FREE Full text] [doi: 10.1111/acem.13325] [Medline: 28960689]
70. Ye T, Xue J, He M, Gu J, Lin H, Xu B, et al. Psychosocial factors affecting artificial intelligence adoption in health care in China: cross-sectional study. J Med Internet Res 2019;21(10):e14316 [FREE Full text] [doi: 10.2196/14316] [Medline: 31625950]

Abbreviations

AI: artificial intelligence
CDSS: clinical decision support system
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Kushniruk; submitted 09.03.21; peer-reviewed by O Asan, D Tao, S Lee; comments to author 13.04.21; revised version received 02.06.21; accepted 07.02.22; published 24.03.22.

Please cite as:
Knop M, Weber S, Mueller M, Niehaves B
Human Factors and Technological Characteristics Influencing the Interaction of Medical Professionals With Artificial Intelligence–Enabled Clinical Decision Support Systems: Literature Review
JMIR Hum Factors 2022;9(1):e28639
URL: https://humanfactors.jmir.org/2022/1/e28639
doi:10.2196/28639
PMID:35323118

Original Paper

Usage and Usability of a National e-Library for Chemotherapy Regimens: Mixed Methods Study

AnnSofie Fyhr[1,2*], MSc; Johanna Persson[2*], PhD; Åsa Ek[2*], PhD

[1]Regional Cancer Centre South, Region Skåne, Lund, Sweden

[2]Ergonomics and Aerosol Technology, Department of Design Sciences, Faculty of Engineering, Lund University, Lund, Sweden

[*]all authors contributed equally

Corresponding Author:
AnnSofie Fyhr, MSc
Regional Cancer Centre South
Region Skåne
Medicon Village
Scheeletorget 1
Lund, SE-223 81
Sweden
Phone: 46 46 275 23 51
Email: ann-sofie.fyhr@skane.se

Abstract

Background: Accurate information about chemotherapy drugs and regimens is needed to reduce chemotherapy errors. A national e-library, as a common knowledge source with standardized chemotherapy nomenclature and content, was developed. Since the information in the library is both complex and extensive, it is central that the users can use the resource as intended.

Objective: The aim of this study was to evaluate the usage and usability of an extensive e-library for chemotherapy regimens developed to reduce medication errors, support the health care staff in their work, and increase patient safety.

Methods: To obtain a comprehensive evaluation, a mixed methods study was performed for a broad view of the usage, including a compilation of subjective views of the users (web survey, spontaneous user feedback, and qualitative interviews), analysis of statistics from the website, and an expert evaluation of the usability of the webpage.

Results: Statistics from the website show an average of just over 2500 visits and 870 unique visitors per month. Most visits took place Mondays to Fridays, but there were 5-10 visits per day on weekends. The web survey, with 292 answers, shows that the visitors were mainly physicians and nurses. Almost 80% (224/292) of respondents searched for regimens and 90% (264/292) found what they were looking for and were satisfied with their visit. The expert evaluation shows that the e-library follows many existing design principles, thus providing some useful improvement suggestions. A total of 86 emails were received in 2020 with user feedback, most of which were from nurses. The main part (78%, 67/86) contained a question, and the rest had discovered errors mainly in some regimen. The interviews reveal that most hospitals use a computerized physician order entry system, and they use the e-library in various ways, import XML files, transfer information, or use it as a reference. One hospital without a system uses the administration schedules from the library.

Conclusions: The user evaluation indicates that the e-library is used in the intended manner and that the users can interact without problems. Users have different needs depending on their profession and their workplace, and these can be supported. The combination of methods applied ensures that the design and content comply with the users' needs and serves as feedback for continuous design and learning. With a broad national usage, the e-library can become a source for organizational and national learning and a source for continuous improvement of cancer care in Sweden.

(JMIR Hum Factors 2022;9(1):e33651) doi:10.2196/33651

KEYWORDS

chemotherapy regimens; user evaluation; standardization; patient safety; chemotherapy; safety; usability; e-library; medication errors

Introduction

Chemotherapy treatments are highly complex, and errors may cause serious harm among patients with cancer, which is a particularly sensitive group owing to impaired tolerance [1,2]. Errors in the chemotherapy process occur at all stages of medication use [2]. However, the prescribing stage plays a key role in the creation of chemotherapy errors [3,4]. Therefore, the use of computerized physician order entry (CPOE) or chemotherapy prescription clinical decision support systems (CDSSs) is recommended [5,6]. Multiple stakeholders, physicians, nurses, and pharmacists, but also patients and their relatives, are involved in the process. This adds to the risk picture, which emphasizes the importance of accurate and timely information about dosages, sequence of therapies, supportive medications, and duration of treatment [7]. Therefore, a common source providing this information was needed. This paper presents a newly developed e-library for chemotherapy regimens with a focus on results from a usage and usability study.

In Sweden, descriptions of chemotherapy regimens were earlier developed and compiled locally within health care organizations. This was mostly because Sweden, with its 10 million inhabitants, is divided into 21 county councils, representing different geographical regions, all having far-reaching autonomy regarding the planning, financing, and operation of the region's health care. In total, there are 17 oncology clinics in Sweden, which manage chemotherapy regimens, of which 7 are located at university hospitals. Additionally, there are several smaller units at other health care clinics, which also manage regimens. The many autonomous clinics have resulted in the same chemotherapy treatments occurring under different names and with different dosages causing uncertainty and risks for mix-ups. To overcome these uncertainties and risks, a national knowledge source for regimens (an e-library) was developed, with standardized nomenclature and content in chemotherapy regimens, also facilitating the exchange of information between hospitals, CPOE systems, and patients.

A standardized national source for chemotherapy regimens can constitute a preventive safety barrier function in the chemotherapy process [8]. Standardization of workflow processes, prescribing, preparation, dispensing, and administration are recommended as safeguards against medication errors [9,10]. Standardized order sets [11], standardized design and architecture [12], and standardized protocols and dosing [13] are beneficial when implementing CPOE systems. In 2020, the HemOnc group (working with a collaborative web-based knowledge platform for oncology professionals in the United States) published a proposal for a standardized nomenclature for chemotherapy regimens that were also compared to the thesaurus of the US National Cancer Institute [14]. In a recently published article, a review of attempts to standardize chemotherapy nomenclature is presented together with recommendations from a European expert panel of oncology pharmacists [15]. In this case, standardization is one way to ensure that all involved health care units have access to a standardized nomenclature and the latest evidence on chemotherapy treatments, which also allows for increased patient safety.

For the national e-library to serve as the intended safety barrier function, it needs to be designed and developed in accordance with user and usability criteria. Shulman et al [16] discusses principles for oncology eHealth records and claims that such systems should be designed to perform logically and straightforwardly, be user-friendly, and always be available to users. The same is applicable for a web-based resource. Based on such an approach the e-library was developed in a user-centered process and has been available since 2015 [17]. Representatives from different user groups were involved throughout the development process, including oncology nurses, physicians, hospital pharmacists, and patients with cancer.

The national e-library [18] contains the following parts: (1) basic facts containing important medical and pharmaceutical information on drugs; (2) regimens presented per diagnosis with an overview (including instructions, precautions, and recommendation for dose reduction), adverse drug reactions (ADRs), and a detailed administration schedule; (3) information sheets for patients per regimen, providing a short description of the treatment and the most common or important ADRs with advice for self-treatment and when to contact the hospital; (4) support documents for health care professionals; and (5) newsletters published after updates of the e-library. Part of a regimen is shown in Figure 1. The primary users are physicians, nurses, and pharmacists. They can access the information in the e-library for reading, printing, or downloading XML files for the CPOE systems used in Sweden. The patients gain access to the information sheets per regimen through their nurse.

Figure 1. Overview of the Cisplatin-Docetaxel regimen for lung cancer. BSA: body surface area, GFR: glomerular filtration rate.

Cisplatin-Docetaxel — Overview — Page 1(2)

Chemotherapy regimen - Lung cancer
Cisplatin-Docetaxel
Indication: Non-small cell lung cancer C34
Cycle intervall: 21 days

Treatment intention: Adjuvant, Curative, Neoadjuvant

Overview

Drug

Substance	Administration	Dilution	Infusion time	Dose/ administration	Dosing	Max dose/ adm.	Max ack. dose
1. Docetaxel	Intravenous infusion	250 mL sodium chloride 0.9% infusion	60 min.	75 mg/m^2	BSA		
2. Cisplatin	Intravenous infusion	1000 mL sodium chloride 0.9% infusion	60 min.	80 mg/m^2	BSA		

Regimen description

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	New cycle day 22
1. Docetaxel	x1																					
2. Cisplatin	x1																					

Emetogenicity: High

Treatment overview: Evaluation after 3-4 cycles.

Instructions for the regimen

Conditions for starting the regimen
Check blood, liver and electrolyte status with renal clearance (Cystatine C, Iohexol, creatinin clearance or equivalent).
Hearing test according to local instructions.
In case of pathological creatinine or if a more favorable side effect profile is desired, switch to Carboplatin-Docetaxel.

Conditions and controls for administration
Cisplatin - Weight or diuresis check.
Docetaxel - Increased preparedness for anaphylactic reaction, greatest risk in cycle 2. Make sure patient has taken premedication.
Check for peripheral neuropathy.

Instructions for prescription
Check blood count including neutrophils, electrolyte status, and creatinine. For start of treatment: neutrophils >1.5, platelets >100 , and leukocytes >2.0.
If S-creatinine above normal value, renal function check with clearance according to local instructions (Cystatine C, Iohexol, creatinine clearance or equivalent). Target value GFR >60.
Docetaxel - Premedication with cortisone, tablet Betamethasone:
Day before treatment give 6 mg Betamethasone, morning and evening.
Day 1 and 2 give 6 mg (total 24 mg) Betamethasone.
The cortisone dose can be reduced in cycle 3 if there is no reaction to the previous treatment.
Cisplatin - during the treatment period at least 4 litres of fluid are given. Intravenous post-hydration can be replaced by liquid.

Dose reduction recommendation
Hematologic toxicities
Nadir for leukocytes <2.0 and/or neutrophils <1.0 - give next cycle with 80% of the doses for both drugs.
If nadir after dose reduction still is leukocytes <2.0 and/or neutrophils <1.0 - reduce doses by an additional 10-15% or change regimen.

Docetaxel - in case of febrile neutropenia or unacceptable adverse drug reactions, consider dose reduction to 75%.

Issued by: regim0fyan
Updated date: 2019-07-08

Version: 1.1
National clinic cancer care guideline group
för Lung cancer

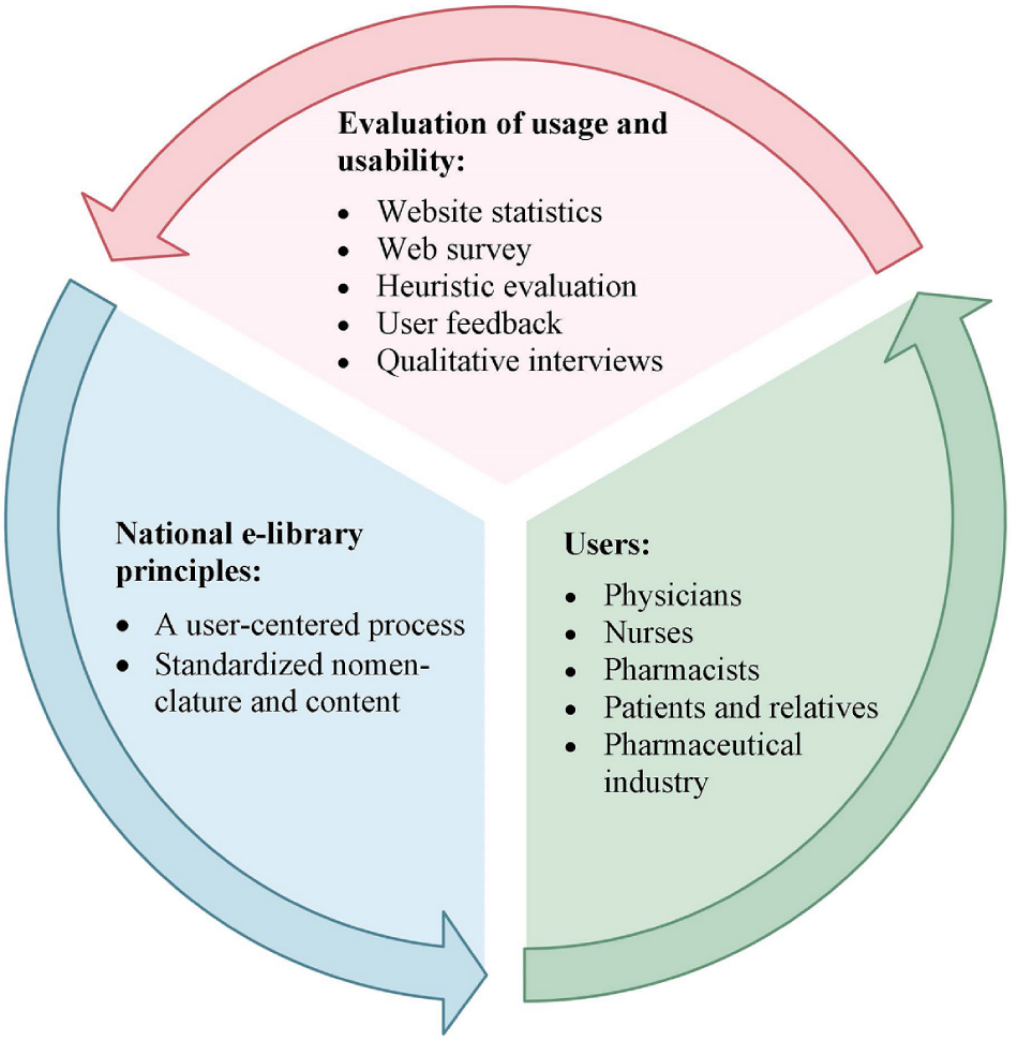
Approved by: regim0chro
Approval date: 2017-06-08

Since the information in the library is both complex and extensive, it is central that the users can utilize the resource as intended. Otherwise, it might not be the source for improved performance and quality in cancer care that is envisioned. The aim of this study was to evaluate the usage and usability of the e-library for chemotherapy regimens developed to reduce medication errors and increase patient safety. This paper presents the results from the evaluation, which will also serve as input for continuous development and improvement.

Methods

A combination of methods was chosen to obtain a comprehensive view of the usage, including a compilation of subjective views of the users, analysis of statistics from the website, and an expert evaluation of the usability of the e-library [19]. The results will help understand where further development is needed or will have the most impact (Figure 2). The evaluation consisted of five parts:

Figure 2. Evaluation of the e-library using multiple methods to conclude its intended usage and usability and to iteratively improve the resource.



1. Website statistics collected from May to December 2020 show the number of users in total, unique users, when they visit the e-library, and from where the visitors were.
2. A web survey for people visiting the e-library was conducted during January 2020. The survey consisted of questions about the visitor's role and function, what information they were looking for, how pleased they were with the visit, and if they had suggestions of improvements.
3. A heuristic evaluation of the user interface [20] was performed in June 2020 to identify details in the user interface design that could be improved to increase usability.
4. Spontaneous user feedback was collected in the form of emails sent to the development team or project leader during 2020, which was obtained from users presenting their role and question or suggestions.
5. Qualitative interviews with 4 nurses, 3 physicians, and 3 pharmacists from various regions in Sweden were performed at the end of 2020 and at the beginning of 2021. As the intention is that the patients should get regimen-related information through their nurse and not through the e-library, no patient interviews were conducted. Interview questions concerned what parts of the e-library they used, how they used it, their experiences of using the e-library in their work compared to how they worked before, and their expectations for further development. The interviews were transcribed verbatim, and a process of transcript analysis was applied to identify themes.

Results

Website Statistics

During the last 8 months of 2020, the average number of visits to the website per month was just above 2500, and the number of unique visitors was 870. Most visitors were from Sweden, with some users from other countries in northern Europe, especially Scandinavia. Visitors from Sweden's 3 largest cities (Stockholm, Gothenburg, and Malmoe) accounted for more than 50% of all visits. Most visits took place on Mondays to Fridays, but there were 5-10 visits per day on weekends.

Web Survey

From the survey with 292 answers, it became clear that the visitors were mainly physicians (49%, 144/292) and nurses (36%, 104/292), and, to some extent, pharmacists (7%, 21/292). The rest were from the pharmaceutical industry and others. Almost 80% (224/292) stated that they searched for regimens and 90% (264/292) stated that they found what they were looking for. Most of the users were satisfied with their visit.

The less satisfied visitors (8%, 22/292) lacked the regimens or diagnostic areas that they were looking for. There were 32 comments, 18 of which asked for more regimens and information sheets for patients. The rest of the comments concerned the design of the survey or were positive reviews.

Heuristic Evaluation

The heuristic evaluation was performed in accordance with the Nielsen and Molich's [20] framework, which implies that a product or system is evaluated in a structured way following ten predefined usability heuristics: (1) visibility of system status; (2) match between systems and the real world; (3) user control and freedom; (4) consistency and standards; (5) error prevention; (6) recognition rather than recall; (7) flexibility and efficiency of use; (8) aesthetic and minimalist design; (9) recognize, diagnose, and recover from errors; and (10) help and documentation. The results show that the e-library follows many of the existing design principles; therefore. it could be concluded to be usable in its current form. Nevertheless, some improvement suggestions were identified, such as moving patient information sheets to a separate tab in the menu, improving the search function by providing suggestions, and improving the start page by moving newsletters to a separate tab.

Spontaneous User Feedback

A total of 86 emails with spontaneous user feedback were received from nurses (50%, 43/86), pharmacists (20%, 17/86), physicians (17%, 15/86), the pharmaceutical industry (8%, 7/86), and others (5%, 4/86). Of the emails, 78% (67/86) concerned questions such as "When will regimens arrive in a certain diagnostic group?" "How long should the drug temozolomide be given before starting radiation therapy?" In 19 of 86 (22%) emails, the sender had discovered errors in the e-library, which mainly concern the regimens, such as patient information sheets where there is a lack of a certain side effect or that the incorrect "treatment intention" had been chosen. All emails were answered, and in cases where there was incorrect or unclear information in the e-library's documents, it was rectified.

Qualitative Interviews

Interview results showed that most hospitals used a CPOE system for their prescription and administration (3 systems exist). One hospital has no system; instead, they use the administration schedule from the e-library. Those with a CPOE system use the library in various ways, including importing the XML files, reading the information in the library and then transferring the information to their CPOE system, or using the library as a reference to check their information. The newsletter is published regularly on the e-library's website, with information on new regimens or basic facts, and detailed information about changes in the regimens was appreciated by most of the interviewees, but not all of them had noticed them. The basic facts are used in various ways depending on profession; the pharmacists focus on preparation, shelf-life, and references to stability studies, the nurses focus on ADRs, and the physicians focus on new drugs. The nurses, especially, were interested in the patient information sheets that are available for most of the regimens. Most hospitals already had their patient information sheets but will switch to the ones in the e-library. The available supporting documents for health care professionals, and especially "Management of side effects associated with immunotherapy with checkpoint inhibitors," are indicated by the nurses and physicians as very useful. Among the expectations for further development was a harmonization of regimens used in different diagnostic areas; for example, the same amount of infusion fluid and infusion times for the same drug.

Discussion

Principal Findings

The results from the user evaluation indicate that the e-library can be concluded to be used in the intended manner, and the users do not have any problems interacting with the knowledge source. From the web surveys and interviews, it becomes clear that it is the content that the users focus on. The usability of the website is not addressed by the users, which may be interpreted as subjective satisfaction in the interaction with the system. The heuristic evaluation showed that there were minor usability issues that should be addressed to improve the overall usability of the website. The web survey showed that the users substantially are satisfied with their visit. Most of the users found what they were looking for, and their main feedback was a desire for more regimens, more diagnostic areas to be covered, and more patient information sheets, which are continuously added and updated. The expressed needs from the users are useful to understand which areas and regimens should be prioritized in the development work. The number of visits indicates that the resource is used extensively, and the geographical span shows that the e-library has emerged as the national resource it is intended to be. Users from almost all Swedish regions exist, although the 3 largest cities stand for the main usage. However, the introduction of the library has proceeded fairly quickly, and it is believed that it is only a matter of time before all regions have adopted the use of the e-library.

The spontaneous user feedback shows clearly that the e-library is used. Contact with the users is vital, and websites generally facilitate a rapid and comprehensive means of knowledge dissemination [7]. Emails from users reporting incorrectness are gratefully received, and they show that errors can slip through despite several checks. The interviews revealed various ways to use the information in the regimens. Some use the XML files to import the regimen to their CPOE system, which is a way to ensure that the transmitted information is accurate. Those transferring the information explained that they had to adjust information in their CPOE system; therefore, is it easier to copy one of their regimens and implement the necessary changes to match the new regimen. Some clinics do not have any CPOE system yet, still relying on handwritten orders and documentation. To transfer information manually is always a risk for introducing errors. Users have different needs depending on whether they are doctors, nurses, or pharmacists and depending on whether they work at a university hospital or a smaller hospital. The e-library can support these different needs from the users.

Automated downloading of the regimens from the e-library to the CPOE systems could improve the process, ensuring updated information without manual handling. However, the CPOE systems do not support this yet, and the process and approval of the updated information at each local clinic must be adapted to the change.

Limitations

This study uses a mix of methods to address the usage and usability of a web-based resource for chemotherapy regimens. One can always discuss which mix of methods is optimal, and the combination used here could certainly be extended to include additional ones. Cognitive walkthrough is one such method, which could have been adopted—it is an explicit and detailed procedure to simulate a uses' problem-solving process at each step through the dialogue with the system having potentially provided additional input in evaluating its usability [21].

The qualitative interviews revealed that nurses have started to replace their old patient information sheets with the ones in the e-library. In a follow-up study, it would be relevant to include patients in the evaluation process.

The study is performed in a geographically delimited system (Sweden), but the article contributes with working methods for how user-driven development contributes to a more standardized working method, which could also provide safer care. That knowledge is useful beyond the Swedish healthcare context. Finally, a longer-term evaluation is required to gain good insight into how the system is used and to evaluate how it contributes to increased safety in cancer care.

Conclusions

The comprehensive user evaluation conducted is an important part of continuing the user-centered process that started already during the development of the e-library. Multiple evaluation methods complement one another by providing input from several perspectives (ie, expert or user, subjective or objective) that may be triangulated and hence identify critical design aspects and user needs [19]. The combination of methods applied in the evaluation presented in this paper included both objective usage statistics, expert methods for usability, and the users' subjective feedback through a web survey and qualitative interviews. This ensures that the design and content of the e-library comply with the users' needs and works as feedback for continuous development and learning [22]. Thus, the evaluation is a vital part assuring that the e-library act as a safety barrier, is well designed and that general design flaw is avoided, a design flaw that otherwise could create new risks in the chemotherapy process. The evaluation will contribute to a deeper understanding of users' judgements about the library content and help to develop strategies for increasing the national usage of the library. With a broad national usage, the e-library can become a source for organizational and national learning and a source for continuous improvement of cancer care in Sweden.

Acknowledgments

We would like to thank the health care personnel who provided feedback on the usage and usability of the e-library. This research received no grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' Contributions

The study was conceived and designed by all authors. AF performed the data collection, as well as most of the data analysis. All authors took part in drafting the manuscript; however, AF performed the main drafting.

Conflicts of Interest

None declared.

References

1. Müller T. Typical medication errors in oncology: analysis and prevention strategies. Onkologie 2003 Dec;26(6):539-544. [doi: 10.1159/000074148] [Medline: 14709927]
2. Weingart SN, Zhang L, Sweeney M, Hassett M. Chemotherapy medication errors. Lancet Oncol 2018 Apr;19(4):e191-e199. [doi: 10.1016/S1470-2045(18)30094-9] [Medline: 29611527]
3. Serrano-Fabiá A, Albert-Marí A, Almenar-Cubells D, Jiménez-Torres NV. Multidisciplinary system for detecting medication errors in antineoplastic chemotherapy. J Oncol Pharm Pract 2010 Jun;16(2):105-112. [doi: 10.1177/1078155209340482] [Medline: 19617304]
4. Fyhr A, Akselsson R. Characteristics of medication errors with parenteral cytotoxic drugs. Eur J Cancer Care (Engl) 2012 Sep;21(5):606-613 [FREE Full text] [doi: 10.1111/j.1365-2354.2012.01331.x] [Medline: 22335508]
5. Kullberg A, Larsen J, Sharp L. 'Why is there another person's name on my infusion bag?' Patient safety in chemotherapy care - a review of the literature. Eur J Oncol Nurs 2013 Apr;17(2):228-235. [doi: 10.1016/j.ejon.2012.07.005] [Medline: 22898657]
6. Rahimi R, Moghaddasi H, Rafsanjani KA, Bahoush G, Kazemi A. Effects of chemotherapy prescription clinical decision-support systems on the chemotherapy process: A systematic review. Int J Med Inform 2019 Feb;122:20-26. [doi: 10.1016/j.ijmedinf.2018.11.004] [Medline: 30623780]

7. Warner JL, Cowan AJ, Hall AC, Yang PC. HemOnc.org: A Collaborative Online Knowledge Platform for Oncology Professionals. J Oncol Pract 2015 May;11(3):e336-e350 [FREE Full text] [doi: 10.1200/JOP.2014.001511] [Medline: 25736385]
8. Hollnagel E. Barriers and Accident Prevention. Farnham: Ashgate; 2004:A.
9. Goldspiel B, Hoffman JM, Griffith NL, Goodin S, DeChristoforo R, Montello CM, et al. ASHP guidelines on preventing medication errors with chemotherapy and biotherapy. Am J Health Syst Pharm 2015 Apr 15;72(8):e6-e35. [doi: 10.2146/sp150001] [Medline: 25825193]
10. Huertas-Fernández MJ, Martínez-Bautista MJ, Rodríguez-Mateos ME, Zarzuela-Ramírez M, Muñoz-Lucero T, Baena-Cañada JM. Implementation of safeguards to improve patient safety in chemotherapy. Clin Transl Oncol 2017 Sep;19(9):1099-1106. [doi: 10.1007/s12094-017-1645-y] [Medline: 28361329]
11. Classen D, Bates DW, Denham CR. Meaningful use of computerized prescriber order entry. J Patient Saf 2010 Mar;6(1):15-23. [doi: 10.1097/PTS.0b013e3181d108db] [Medline: 22130298]
12. Kukreti V, Cosby R, Cheung A, Lankshear S, ST Computerized Prescriber Order Entry Guideline Development Group. Computerized prescriber order entry in the outpatient oncology setting: from evidence to meaningful use. Curr Oncol 2014 Aug;21(4):e604-e612 [FREE Full text] [doi: 10.3747/co.21.1990] [Medline: 25089110]
13. Dabliz R, Poon SK, Ritchie A, Burke R, Penm J. Usability evaluation of an integrated electronic medication management system implemented in an oncology setting using the unified theory of the acceptance and use of technology. BMC Med Inform Decis Mak 2021 Jan 06;21(1):4 [FREE Full text] [doi: 10.1186/s12911-020-01348-y] [Medline: 33407411]
14. Rubinstein SM, Yang PC, Cowan AJ, Warner JL. Standardizing Chemotherapy Regimen Nomenclature: A Proposal and Evaluation of the HemOnc and National Cancer Institute Thesaurus Regimen Content. JCO Clin Cancer Inform 2020 Jan;4:60-70 [FREE Full text] [doi: 10.1200/CCI.19.00122] [Medline: 31990580]
15. Terkola R, Bardin C, Lizeaga Cundin G, Zeinab N, Crul M. Identifying options for oncology therapy regimen codification to improve standardization-combined results of an expert panel and a review. J Clin Pharm Ther 2021 Oct;46(5):1238-1244 [FREE Full text] [doi: 10.1111/jcpt.13402] [Medline: 33687085]
16. Shulman LN, Miller RS, Ambinder EP, Yu PP, Cox JV. Principles of Safe Practice Using an Oncology EHR System for Chemotherapy Ordering, Preparation, and Administration, Part 1 of 2. J Oncol Pract 2008 Jul;4(4):203-206 [FREE Full text] [doi: 10.1200/JOP.0847501] [Medline: 29447501]
17. Fyhr A, Borell J, Jerkeman M, Ek Å. National e-library for standardized chemotherapy regimens. Acta Oncol 2020 Sep;59(9):1079-1083. [doi: 10.1080/0284186X.2020.1764097] [Medline: 32401129]
18. Regionala Cancercentrum i Samverkan. URL: https://kunskapsbanken.cancercentrum.se/ [accessed 2022-02-10]
19. Price M, Weber J, Bellwood P, Diemert S, Habibi R. Evaluation of eHealth System Usability and Safety. In: Lau F, Kuziemsky C, editors. Handbook of eHealth Evaluation: An Evidence-based Approach. Victoria, BC: University of Victoria; Feb 27, 2017.
20. Nielsen J, Molich R. Heuristic evaluation of user interfaces. 1990 Presented at: CHI90: Conference on Human Factors in Computing; April 1-5, 1990; Seattle, WA p. 249-256. [doi: 10.1145/97243.97281]
21. Nielsen J, Mack RL. Usability Inspection Methods. New York, NY: Wiley; May 09, 1994.
22. Carayon P, Salwei ME. Moving toward a sociotechnical systems approach to continuous health information technology design: the path forward for improving electronic health record usability and reducing clinician burnout. J Am Med Inform Assoc 2021 Apr 23;28(5):1026-1028 [FREE Full text] [doi: 10.1093/jamia/ocab002] [Medline: 33537756]

Abbreviations

ADR: adverse drug reaction
CDSS: clinical decision support system
CPOE: computerized physician order entry

Edited by A Kushniruk; submitted 17.09.21; peer-reviewed by H Moghaddasi, M Hellaby; comments to author 28.11.21; revised version received 09.12.21; accepted 19.12.21; published 17.02.22.

Please cite as:
Fyhr A, Persson J, Ek Å
Usage and Usability of a National e-Library for Chemotherapy Regimens: Mixed Methods Study
JMIR Hum Factors 2022;9(1):e33651
URL: https://humanfactors.jmir.org/2022/1/e33651
doi:10.2196/33651
PMID:35175199

Original Paper

Usability of Wearable Multiparameter Technology to Continuously Monitor Free-Living Vital Signs in People Living With Chronic Obstructive Pulmonary Disease: Prospective Observational Study

Grace Hawthorne[1], PhD; Neil Greening[1,2], MD, PhD; Dale Esliger[3], PhD; Samuel Briggs-Price[1], MA; Matthew Richardson[2], PhD; Emma Chaplin[1], BSc; Lisa Clinch[1], BSc; Michael C Steiner[1,2], MD, PhD; Sally J Singh[1,2], PhD; Mark W Orme[1,2], PhD

[1]Centre for Exercise and Rehabilitation Science, National Institute for Health Research Leicester Biomedical Research Centre – Respiratory, University Hospitals of Leicester National Health Service Trust, Leicester, United Kingdom

[2]Department of Respiratory Sciences, University of Leicester, Leicester, United Kingdom

[3]School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, United Kingdom

Corresponding Author:
Grace Hawthorne, PhD
Centre for Exercise and Rehabilitation Science
National Institute for Health Research Leicester Biomedical Research Centre – Respiratory
University Hospitals of Leicester National Health Service Trust
Groby Road
Leicester, LE3 9QP
United Kingdom
Phone: 44 116258 ext 3035
Email: grace.france1@nhs.net

Abstract

Background: Vital signs monitoring (VSM) is routine for inpatients, but monitoring during free-living conditions is largely untested in chronic obstructive pulmonary disease (COPD).

Objective: This study investigated the usability and acceptability of continuous VSM for people with COPD using wearable multiparameter technology.

Methods: In total, 50 people following hospitalization for an acute exacerbation of COPD (AECOPD) and 50 people with stable COPD symptoms were asked to wear an Equivital LifeMonitor during waking hours for 6 weeks (42 days). The device recorded heart rate (HR), respiratory rate (RR), skin temperature, and physical activity. Adherence was defined by the number of days the vest was worn and daily wear time. Signal quality was examined, with thresholds of ≥85% for HR and ≥80% for RR, based on the device's proprietary confidence algorithm. Data quality was calculated as the percentage of wear time with acceptable signal quality. Participant feedback was assessed during follow-up phone calls.

Results: In total, 84% of participants provided data, with average daily wear time of 11.8 (SD 2.2) hours for 32 (SD 11) days (average of study duration 76%, SD 26%). There was greater adherence in the stable group than in the post-AECOPD group (≥5 weeks wear: 71.4% vs 45.7%; *P*=.02). For all 84 participants, the median HR signal quality was 90% (IQR 80%-94%) and the median RR signal quality was 93% (IQR 92%-95%). The median HR data quality was 81% (IQR 58%-91%), and the median RR data quality was 85% (IQR 77%-91%). Stable group BMI was associated with HR signal quality (r_s=0.45, *P*=.008) and HR data quality (r_s=0.44, *P*=.008). For the AECOPD group, RR data quality was associated with waist circumference and BMI (r_s=–0.49, *P*=.009; r_s=–0.44, *P*=.02). In total, 36 (74%) participants in the Stable group and 21 (60%) participants in the AECOPD group accepted the technology, but 10 participants (12%) expressed concerns with wearing a device around their chest.

Conclusions: This wearable multiparametric technology showed good user acceptance and was able to measure vital signs in a COPD population. Data quality was generally high but was influenced by body composition. Overall, it was feasible to continuously measure vital signs during free-living conditions in people with COPD symptoms but with additional challenges in the post-AECOPD context.

(JMIR Hum Factors 2022;9(1):e30091) doi:10.2196/30091

KEYWORDS

chronic obstructive pulmonary disease; digital health; physical activity; respiratory rate; wearable technology; wearable device; vital signs monitor

Introduction

Chronic obstructive pulmonary disease (COPD) is the third leading cause of death worldwide [1]. People living with COPD may experience an acute exacerbation (ie, acute exacerbation of COPD [AECOPD]), which reduces their quality of life and increases the risk of premature mortality [2].

While often defined by worsening of respiratory symptoms, an AECOPD is associated with changes in heart rate, oxygen saturation [3], and respiratory rate [4], with such vital signs monitored routinely as an inpatient. Transferring this monitoring to the daily lives of outpatients has been challenging but, by doing so, it may be possible to recognize deterioration in health [5,6]. Studies have remotely monitored symptoms through pulse oximetry or spirometry to identify changes in patient health for a while now [7-12], as patients typically find it difficult to identify small day-to-day variations in symptoms [13,14]. The use of remote patient monitoring following hospitalization for an AECOPD is less common [15-18], despite this population being at high risk of readmission to hospital [19,20].

When deploying technological solutions, patient burden is an important barrier to success. To date, studies have relied on patients actively taking daily measurements, such as from pulse oximeters [7,9,10,12,15-17,21,22]. Patient-driven measurements could result in recall bias, errors in data collection [23], and reduced compliance [24-26]. Other studies have required patients to use multiple devices to measure vital signs [12,26,27], which adds to patient burden, with the additional complication of managing multiple devices leading to reduced adherence [26]. It can be even more challenging for individuals following hospitalization for an AECOPD to engage with digital health technologies [17,28,29], perhaps owing to greater disease severity [30]. Providing that it is comfortable and accepted by patients, wearable technology could facilitate free-living health monitoring.

Accordingly, we aimed to determine whether (1) we can measure vital signs using a novel wearable device post hospitalization for AECOPD and during the stable phase of COPD, (2) there are patient characteristics associated with adherence and data quality, and (3) measures of feasibility are different between people post AECOPD and those with stable COPD symptoms.

Methods

Recruitment

We performed a prospective, observational cohort study of people living with COPD admitted to hospital for an AECOPD (AECOPD group) and people with stable COPD symptoms (Stable group). This single-center study was undertaken between January 2018 and December 2019 at the University Hospitals of Leicester, the United Kingdom, where individuals were recruited from hospital wards or the pulmonary rehabilitation (PR) service.

People with an AECOPD were screened by COPD specialist nurses and recruited when medically stable and close to being discharged. People with stable disease were screened by the PR team at their initial PR assessment and enrolled prior to starting their PR program. Inclusion criteria were as follows: being ≥18 years old; having a confirmed clinical diagnosis of COPD from spirometry data in medical records; and for the AECOPD group, an admission with a primary diagnosis of exacerbation of COPD was required. Participants were excluded if they had a physical or visual impairment or comorbidities that prevent participation, required palliative care, were participating in another study, or were unable or unwilling to provide written informed consent.

Vital Signs Measurements Using an Equivital LifeMonitor

All participants were asked to wear an Equivital EQ02+ LifeMonitor device (Equivital) (hereby, “vest”) during waking hours for 6 consecutive weeks. During their baseline visit, participants practiced putting on and removing the vest with a researcher first, and then independently while supervised, and they were also given written and visual instructions to take home. Participants were asked to remove the vest during water-based activities and to charge the sensor electronics module (SEM) overnight, at least every other night.

Our patient and public involvement (PPI) group contributed to the design of the study. Specifically, members selected the Equivital LifeMonitor, from a choice of three wearable devices, and provided feedback on the duration of wear period. They also provided feedback on study documentation, wording, and verbal description of the study for recruitment purposes.

The AECOPD group were asked to start wearing the vest following discharge from hospital, making the day after discharge the first day of wear. The Stable group were asked to wear the vest following their baseline study assessment (after initial PR assessment). The day after the baseline visit was the first day of wear (Multimedia Appendix 1).

Participants were contacted by a researcher via telephone 1-3 days after their baseline visit to evaluate acceptability of the vest. Participants had further follow-up telephone calls on a fortnightly basis for troubleshooting purposes or could contact the study team on an *ad hoc* basis.

The vest measured heart rate (HR), respiratory rate (RR), skin temperature (ST), and physical activity (PA) [31]. PA was classified as stationary or ambulatory using an inbuilt triaxial accelerometer. HR was obtained using built-in ECG electrodes, RR was recorded with a built-in expansion belt, and ST was measured using a thermometer in the SEM (15-second epoch).

Measures of Feasibility

Definitions and criteria for feasibility indicators are specified in Table 1. Adherence was defined by the number of days the

vest was worn during the 6-week study period and the daily wear time. Missing data were examined and classified as either battery depletion (failing to charge the SEM) or nonwear.

The signal quality of HR and RR were based on the proprietary confidence algorithm, accounting for activity and connection artefacts. Based on manufacturer's recommendations, a signal quality threshold of 85% was used for HR and 80% for RR to indicate whether each 15-second value was deemed acceptable. Data quality for HR and RR was defined as the percentage of daily wear time with acceptable signal quality. Signal quality and data quality were used to identify the confidence of the data generated by the vest.

Field notes from the follow-up phone calls were analyzed to ascertain any common problems that participants had with the technology or other aspects of the study participation. Acceptability was defined as reporting no problems with the technology.

Table 1. Measures of feasibility.

Measure (unit)	Definition	Calculation
Duration worn (days)	Number of days that a participant wore the vest across 6 weeks (42 days)	Number of days (out of 42) deemed as worn with minimum wear time thresholds (1-16 hours)
Wear time (hours)	The duration for which the vest was worn in a single day	Sum of the time with a heart rate of >25 beats/min and a skin temperature of >25°C
Heart rate signal quality (%)	The confidence that the heart rate data obtained are accurate	Average heart rate confidence when the vest was worn
Respiratory rate signal quality (%)	The confidence that the respiratory rate data obtained are accurate	Average respiratory rate confidence when the vest was worn
Heart rate data quality (%)	The proportion of daily wear time when the heart rate signal quality was ≥85%	Proportion of time that the heart rate confidence was ≥85% when the vest was worn
Respiratory rate data quality (%)	The proportion of daily wear time when the respiratory rate signal quality was ≥80%	Proportion of time that the respiratory rate confidence was ≥80% when the vest was worn
Skin temperature data quality (%)	The proportion of daily wear time when respiratory rate signal quality or heart rate signal quality was valid	Proportion of time that the respiratory rate confidence was ≥80% or the heart rate confidence was ≥85% when the vest was worn

Vital Sign Measurements

Vital signs examined were HR, RR, ST, and PA. HR and RR were calculated as the average HR and RR, respectively, during wear time. PA was calculated as the proportion of daily wear time when the patient was ambulatory.

Participant Characteristics

Demographics, clinical histories, comorbidities, and spirometry data were obtained from medical records or information provided by participants. Height and weight were obtained from medical records or measured. Chest circumference and waist circumference were measured.

The Medical Research Council dyspnea scale [32] was used to measure breathlessness.

Statistical Analyses

No formal sample size calculation was undertaken for this feasibility study. A sample size of 50 participants per group was decided on the basis of potential suitable participants, logistics, and resources available.

Data were analyzed using R (version 4.0.0). Continuous variables distributions were tested for normality. Data are reported as mean (SD) or median (IQR) and differences between groups were assessed using a 2-sample unpaired *t* test or Mann–Whitney *U* test, respectively. Frequency comparisons between groups were assessed using the Fisher test. The Spearman rank correlation coefficient (r_s) was used to analyze associations between variables (Cronbach α=.05).

Ethics Approval

All participants provided written informed consent (Research Ethics Committee 15/LO/2055) and the study was prospectively registered (ISRCTN12855961).

Results

Recruitment and Participant Characteristics

Figure 1 outlines recruitment details, reasons for withdrawal, and completion rate for the AECOPD and Stable groups. The AECOPD group had a lower BMI, more severe dyspnea, more hospital admissions, and more frequent exacerbations, but they were otherwise similar to the Stable group (Table 2).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flowchart for the AECOPD and Stable groups. AECOPD: acute exacerbation of chronic obstructive pulmonary disease; COPD: chronic obstructive pulmonary disease; DNA: did not attend; PIS: patient information sheet; PR: pulmonary rehabilitation.

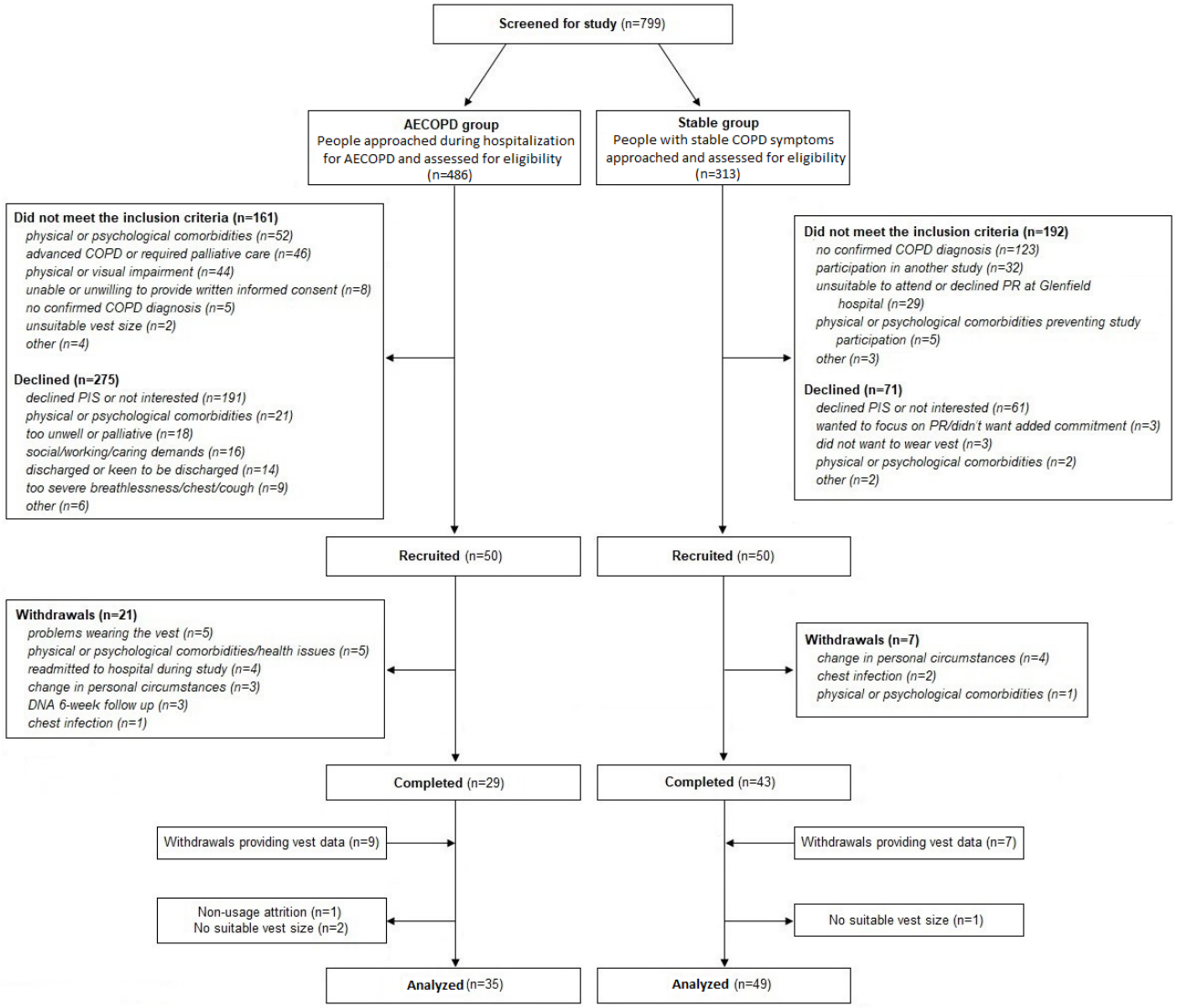


Table 2. Baseline characteristics of the AECOPD[a] and Stable groups.

Characteristics	AECOPD group (n=35)	Stable group (n=49)	*P* value
Male, n (%)	19 (54.3)	27 (55.1)	>.99
Age (years), mean (SD)	67.6 (9.7)	66.7 (9.2)	.77
BMI (kg/m^2), mean (SD)	25.2 (6.0)	28.0 (6.7)	.04
Chest circumference (inches), mean (SD)	37.9 (4.4)	39.0 (4.0)	.29
Waist circumference (cm), mean (SD)	94.0 (17.1)	97.4 (15.2)	.39
Forced expiratory volume in 1 second (% predicted), mean (SD)	44.8 (22.3)	53.5 (27.4)	.16
Forced expiratory volume in 1 second–forced vital capacity ratio, mean (SD)	0.43 (0.16)	0.50 (0.16)	.14
Medical Research Council dyspnea grade, n (%)			
2	8 (22.9)	15 (30.6)	.47
3	5 (14.3)	17 (34.7)	.045
4	14 (40.0)	16 (32.7)	.499
5	8 (23.8)	1 (2.0)	.003
Smoking status, n (%)			
Never	0	3 (7.1)[b]	.81
Ex-smoker	23 (65.7)	29 (69.1)[b]	.33
Current	12 (34.3)	10 (23.8)[b]	.25
Pack years (years), median (IQR)	48.0 (35.5-63.8)	40.0 (27.0-50.0)[c]	.09
Oxygen use, n (%)	4 (11.4)	3 (7.3)[d]	.70
Hospital admissions in the last 12 months, median (IQR)	1.5 (1.0-2.8)	0 (0-1.0)[e]	<.001
Exacerbations in the last 12 months, median (IQR)	3.0 (2.0-4.0)	0.5 (0-3.0)[e]	.009
Physical activity (hours/day)[f], median (IQR)	1.3 (0.9-1.7)	1.5 (1.1-2.0)	.03

[a]AECOPD: acute exacerbation of chronic obstructive pulmonary disease.

[b]Missing data (n=7).

[c]Missing data (n=23).

[d]Missing data (n=8).

[e]Missing data (n=25).

[f]Calculated from the Equivital LifeMonitor.

Feasibility Measures

For all 84 participants, the vest was worn for a median of 37.0 (IQR 27.8-40.0) days and the median daily wear time was 12.0 (IQR 10.8-13.1) hours. The median HR signal quality was 90% (IQR 80%-94%), and the median RR signal quality was 93% (IQR 92%-95%; Figure 2A). The median HR data quality was 81% (IQR 58%-91%), and the median RR data quality was 85% (IQR 77%-91%; Figure 2B).

There were no significant between-group differences in the number of days the vest was worn, the longest number of consecutive days worn, or the average daily wear time (Table 3). The AECOPD group spent a significantly lesser median time ambulatory (10.1%, IQR 8.6%-15.0% vs 13.4%, IQR 9.5%-19.3%; *P*=.03) and showed a lower median HR signal quality (88.5%, IQR 75.8%-92.6% vs 92.3%, IQR 81.0%-96.4%; *P*=.04) than the Stable group.

Figure 2. HR and RR (A) signal quality and (B) data quality. Data are shown as box plots composed of the 25th percentile (lower extremity of the box), the median (central line of the box), and the 75th percentile (upper extremity of the box). The lines outside each box correspond to the minimum and maximum values. HR: heart rate; RR: respiratory rate.

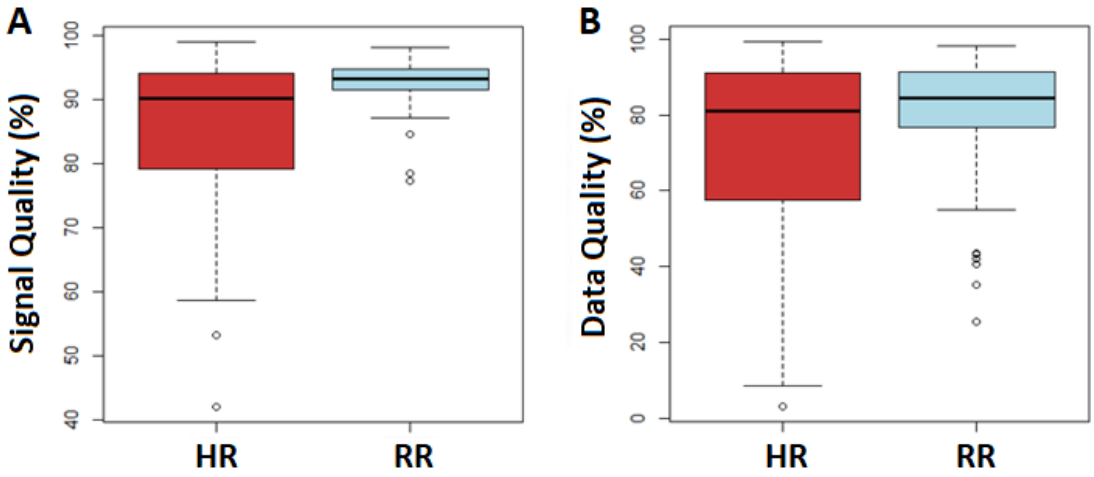


Table 3. Feasibility and vital sign measures for the AECOPD[a] and Stable groups.

Measures	AECOPD group (n=35)	Stable group (n=49)	*P* value
Duration worn (days; maximum, 42 days), median (IQR)	33.0 (23.0-39.5)	38.0 (32.0-40.0)	.11
Duration worn (weeks), n (%)			
>5	16 (45.7)	35 (71.4)	.02
4-5	8 (22.9)	4 (8.2)	.11
3-4	3 (8.6)	5 (10.2)	>.99
2-3	2 (5.7)	3 (6.1)	>.99
1-2	4 (11.4)	1 (2.1)	.16
<1	2 (5.7)	1 (2.1)	.57
Longest number of consecutive days worn (days), median (IQR)	17.0 (7.0-29.5)	21.0 (10.0-38.0)	.34
Occasions of missing days (days), median (IQR)	2.0 (0.25-5.75)	3.0 (1.0-5.0)	.56
Flat sensor electronics module battery depletion during wear, n (%)	0 (0%)	3 (6%)	.26
Wear time (hours), median (IQR)	12.1 (10.4-12.6)	11.9 (10.9-13.3)	.36
Heart rate signal quality (%), median (IQR)	88.5 (75.8-92.6)	92.3 (81.0-96.4)	.04
Respiratory rate signal quality (%), median (IQR)	93.4 (91.5-94.9)	93.2 (91.8-94.6)	.61
Heart rate data quality (%), median (IQR)	78.1 (55.7-88.0)	86.0 (58.8-95.5)	.10
Respiratory rate data quality (%), median (IQR)	83.7 (75.2-91.5)	85.5 (77.2-89.9)	.97
Skin temperature data quality (%), median (IQR)	95.6 (89.1-97.9)	97.0 (92.1-98.9)	.27
Heart rate (beats/min), mean (SD)	84.4 (10.3)	84.4 (10.2)	.97
Respiratory rate (breaths/min), mean (SD)	20.6 (3.5)	20.3 (3.2)	.71
Skin temperature (°C), mean (SD)	34.2 (0.81)	34.3 (0.98)	.90
Stationary (hours), median (IQR)	10.2 (9.7-11.4)	10.2 (9.5-11.5)	.88
Physical activity (% of wear time), median (IQR)	10.1 (8.6-15.0)	13.4 (9.5-19.3)	.03

[a]AECOPD: acute exacerbation of chronic obstructive pulmonary disease.

Participant Acceptability

From follow-up phone calls, 21 participants (60%) in the AECOPD group and 36 participants (74%) in the Stable group found the vest acceptable. Five (14%) participants in the AECOPD group reported that they did not wear the vest while unsettled or feeling unwell after returning home. Five (10%) participants in the Stable group and 5 (14%) participants in the AECOPD group experienced some discomfort wearing the vest. Three (9%) participants and 5 (10%) participants in the

AECOPD and Stable groups, respectively, reported that they did not wear the vest on days on which they felt unwell. Four (11%) participants in the AECOPD group had problems removing the SEM from the cradle of the vest, compared to one participant (2%) in the Stable group.

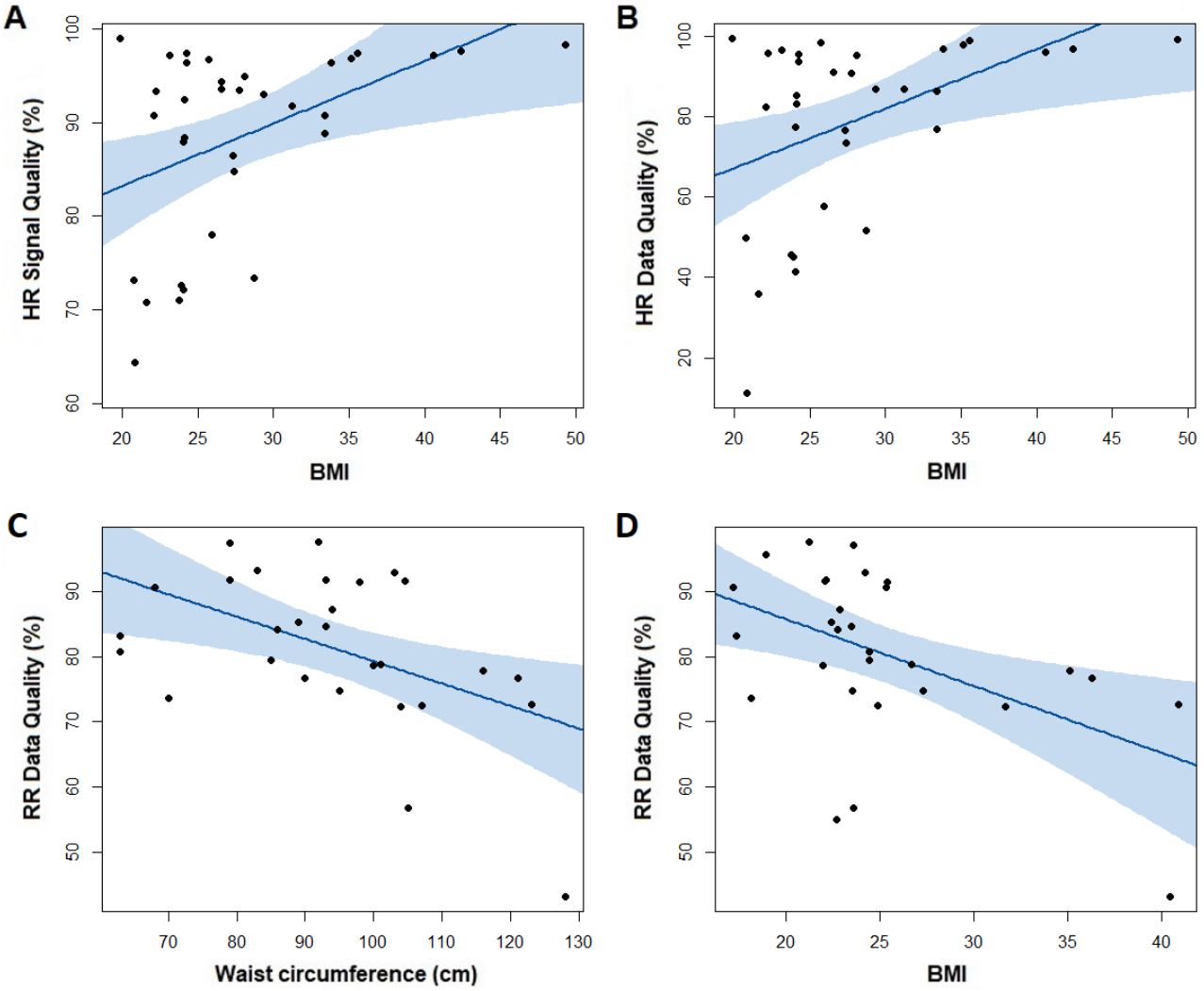
Vest Fitting

For the whole sample, 33 (40%) participants were allocated a larger vest size, which was comparable between groups. Compared to those who completed the study, a larger proportion of participants who withdrew required a larger vest size than the manufacturer's guidance (71% vs 34%, *P*=.005). There were no associations between participants vest fitting and other feasibility measures (Multimedia Appendix 1).

Relationships Between Feasibility Measures and Anthropometric Characteristics

For the Stable group, HR signal quality and HR data quality were positively correlated with BMI (r_s=0.45, *P*=.008; r_s=0.44, *P*=.008; Figures 3A and 3B). For the AECOPD group, RR data quality was negatively correlated with waist circumference and BMI (r_s=–0.49, *P*=.009; r_s=–0.44, *P*=.02; Figures 3C and 3D).

Figure 3. Associations within the Stable group (n=49) between (A) HR signal quality and BMI and (B) HR data quality and BMI, and associations within the AECOPD group (n=35) between (C) RR data quality and waist circumference and (D) RR data quality and BMI. The shaded area indicates 95% CI and *P* values and correlations calculated using Spearman ρ. AECOPD: acute exacerbation of chronic obstructive pulmonary disease; HR: heart rate; RR: respiratory rate.



Discussion

Principal Findings

In this study, it was possible to continuously measure vital signs (RR, HR, ST, and PA) in free-living conditions using multiparameter wearable technology for people living with COPD after hospitalization and during stable symptoms. Measurement of vital signs using our technology was more challenging post AECOPD. The Equivital LifeMonitor was acceptable to some people with stable COPD symptoms and post AECOPD and produced data of sufficient quality. It should be noted that some participants felt uncomfortable with a device around their chest, and data quality was influenced by body composition. Overall, continuous VSM during daily life is possible, and its potential utility for supporting patient post AECOPD should be explored.

In this study, some participants stated that they did not want to wear a device around their chest. As well as the physical implications of a vest-like device, people with COPD may be influenced by the psychological impact of a device around the

chest, as reported previously [26]. While most participants in our study reported no discomfort wearing the vest, similar to a previous study measuring respiratory rate with a chest belt [26], some participants reported that the vest felt restrictive at times or made them feel breathless. Perceived breathlessness could have influenced vest acceptability, as COPD populations tend to prefer watch-like devices [33]. However, measuring respiratory rate from such devices is challenging [34]. The use of a wearable, wireless patch has been successful in continuously monitoring vital signs of inpatients [35], and although this would reduce the perceived breathlessness with a vest-like device, Rubio et al [34] reported that a chest band could measure the RR more reliably than patch-like devices.

Recruitment was more challenging in the AECOPD group, with 15.4% of eligible patients recruited for the AECOPD group, compared to 41.3% in the Stable group, as seen in previous similar studies [9,17]. In a study where patients were asked to record oxygen saturation (SpO_2) and RR using 2 separate devices for 2 months, 79% measured SpO_2 and 60% measured RR three times per day, while 98% and 83% measured SpO_2 and RR, respectively, once per day [26]. In this study, we observed a comparable adherence to findings when patients took daily measures (84% of participants provided vest data) but were able to capture a vast amount of data (on average 12 hours of data). Compared to the AECOPD group, a greater proportion of the Stable group wore the vest for >5 weeks (46% vs 71%). Technology continues to develop multiparametric wearable devices [36-39] to reduce patient burden, but it is possible that this remains a significant barrier in an AECOPD population with lower digital literacy.

Participant feedback from telephone calls suggested that the vest was acceptable overall, with 68% of participants reporting no problems. Some participants reported that they chose not to wear the vest on the days that they felt unwell. Vitacca et al [18] asked participants to complete a weekly 12-item Respicard (recording symptoms, SpO_2, and HR) for 6 months and identified that participants with worse respiratory values had poorer adherence. In this study, 10% of the AECOPD group withdrew because they experienced problems wearing the vest, and 14% of participants in the AECOPD group struggled to engage with the device once returning home. Following discharge from hospital post AECOPD, symptoms remain elevated, and it takes time for patients to recover to their normal symptoms and daily activities [7,40]. Despite our single piece of technology reducing the need for patients to measure multiple vital signs and the observational nature of the study, the greater symptom burden in an AECOPD population reduced adherence to wearable technology.

Our results show that some patients were unable to participate as their chest size exceeded the maximum vest size, or they felt that the maximum vest size was not a suitable fitting. The Equivital LifeMonitor used in this study was originally designed to monitor vital signs in a military population [41]. A greater proportion of those who withdrew required a larger vest size than the completers (71% vs 34%). Existing wearable technology is more broadly marketed toward a healthy population and is typically not tailored for people living with COPD. While the form of the technology used in this study was generally acceptable, advancements in more discrete technologies are needed.

Similar to previous reports in healthy men [31], this study shows that HR and RR measurements obtained from the vest are of sufficient quality in a COPD population. Evidence suggests accurate vital signs measurements of clothing monitors such as the Zephyr BioHarness and Hexoskin [42,43], but our PPI members found such devices challenging to put on and remove owing to their tight-fitting nature. Despite the technology used in this study being tested by our PPI group, physical impairments affected participants' ability to wear the vest and charge the SEM. Compared to 2% of the Stable group, 11% of the AECOPD group reported problems putting the SEM in the cradle, with some needing help from a cohabitant. The HR signal and data quality were also worse for people with a lower BMI, which is seen more often in an AECOPD population [44-46]. This may be owing to lower body composition and lower conductance [47]; therefore, a weaker connection between the skin and the electrodes embedded in the vest. These problems have been observed elsewhere [26] and must be considered by manufacturers, researchers, and clinicians when selecting digital health technologies.

Limitations

Although the number of patients assessed in this study was low as a proportion of patients screened, introducing the possibility of selection bias, recruitment is often challenging post AECOPD [17,28,29]. Our single piece of technology aimed to passively capture multiple vital signs; however, some participants may prefer active participation to obtain recordings. It was not possible to measure SpO_2 and blood pressure in the continuous and unobtrusive manner in line with this study. The lack of an age-matched healthy control group prevented us from identifying the unique difficulties with the use of multiparameter technology in a COPD population. This study may have benefitted from measuring vital signs overnight, to obtain individualized "baseline" vital sign values. A more rigorous qualitative exploration of participants' experiences would have provided greater insights than telephone call field notes.

Conclusions

Following hospitalization for AECOPD and during stable symptoms, it was possible to continuously measure RR, HR, ST, and PA using multiparameter wearable technology during free-living conditions. The Equivital LifeMonitor was acceptable to participants and produced data of sufficient quality, despite some reports of discomfort with wearing a device around the chest and data quality influenced by body composition. Overall, continuous VSM during daily life is possible for people living with COPD and its potential utility for supporting patients post AECOPD should be further explored.

Acknowledgments

The authors would like to thank the University Hospitals of Leicester Pulmonary Rehabilitation team, the chronic obstructive pulmonary disease (COPD) Specialist Nursing team, and the Centre for Exercise and Rehabilitation Science, National Institute for Health Research (NIHR) Leicester Biomedical Research Centre – Respiratory, for their contribution towards data collection. The authors acknowledge support from the NIHR Leicester Biomedical Research Centre, which is a partnership between the University Hospitals of Leicester National Health Service (NHS) Trust and the University of Leicester, and acknowledge the support from the NIHR Collaboration for Leadership in Applied Health Research and Care East Midlands. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, or the Department of Health. The authors disclose receipt of financial support from the Pfizer OPEN AIR scheme (grant WP1462736). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures.
[DOCX File , 240 KB - humanfactors_v9i1e30091_app1.docx]

References

1. The top 10 causes of death. World Health Organization. 2020 Dec 09. URL: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [accessed 2021-01-28]
2. Soler-Cataluña JJ, Martínez-García MA, Román Sánchez P, Salcedo E, Navarro M, Ochando R. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. Thorax 2005 Nov;60(11):925-931 [FREE Full text] [doi: 10.1136/thx.2005.040527] [Medline: 16055622]
3. Hurst JR, Donaldson GC, Quint JK, Goldring JJ, Patel AR, Wedzicha JA. Domiciliary pulse-oximetry at exacerbation of chronic obstructive pulmonary disease: prospective pilot study. BMC Pulm Med 2010 Oct 20;10:52 [FREE Full text] [doi: 10.1186/1471-2466-10-52] [Medline: 20961450]
4. Yañez AM, Guerrero D, Pérez de Alejo R, Garcia-Rio F, Alvarez-Sala JL, Calle-Rubio M, et al. Monitoring breathing rate at home allows early identification of COPD exacerbations. Chest 2012 Dec;142(6):1524-1529. [doi: 10.1378/chest.11-2728] [Medline: 22797131]
5. Al Rajeh A, Hurst J. Monitoring of physiological parameters to predict exacerbations of chronic obstructive pulmonary disease (COPD): a systematic review. J Clin Med 2016 Nov 25;5(12):108 [FREE Full text] [doi: 10.3390/jcm5120108] [Medline: 27897995]
6. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote patient monitoring via non-invasive digital technologies: a systematic review. Telemed J E Health 2017 Jan;23(1):3-17 [FREE Full text] [doi: 10.1089/tmj.2016.0051] [Medline: 27116181]
7. Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2000 May;161(5):1608-1613. [doi: 10.1164/ajrccm.161.5.9908022] [Medline: 10806163]
8. Pedone C, Chiurco D, Scarlata S, Incalzi RA. Efficacy of multiparametric telemonitoring on respiratory outcomes in elderly people with COPD: a randomized controlled trial. BMC Health Serv Res 2013 Mar 06;13:82 [FREE Full text] [doi: 10.1186/1472-6963-13-82] [Medline: 23497109]
9. Cordova FC, Ciccolella D, Grabianowski C, Gaughan J, Brennan K, Goldstein F, et al. A telemedicine-based intervention reduces the frequency and severity of COPD exacerbation symptoms: a randomized, controlled trial. Telemed J E Health 2016 Feb;22(2):114-122 [FREE Full text] [doi: 10.1089/tmj.2015.0035] [Medline: 26259074]
10. Sund Z, Powell T, Greenwood R, Jarad N. Remote daily real-time monitoring in patients with COPD --a feasibility study using a novel device. Respir Med 2009 Sep;103(9):1320-1328 [FREE Full text] [doi: 10.1016/j.rmed.2009.03.017] [Medline: 19375294]
11. Pitta F, Troosters T, Spruit MA, Probst VS, Decramer M, Gosselink R. Characteristics of physical activities in daily life in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2005 May 01;171(9):972-977. [doi: 10.1164/rccm.200407-855OC] [Medline: 15665324]
12. Miłkowska-Dymanowska J, Białas AJ, Obrębski W, Górski P, Piotrowski WJ. A pilot study of daily telemonitoring to predict acute exacerbation in chronic obstructive pulmonary disease. Int J Med Inform 2018 Aug;116:46-51 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.04.013] [Medline: 29887234]
13. Burton C, Pinnock H, McKinstry B. Changes in telemonitored physiological variables and symptoms prior to exacerbations of chronic obstructive pulmonary disease. J Telemed Telecare 2015 Jan;21(1):29-36. [doi: 10.1177/1357633X14562733] [Medline: 25475218]

14. Calverley P, Pauwels Dagger R, Löfdahl CG, Svensson K, Higenbottam T, Carlsson L, et al. Relationship between respiratory symptoms and medical treatment in exacerbations of COPD. Eur Respir J 2005 Sep;26(3):406-413 [FREE Full text] [doi: 10.1183/09031936.05.00143404] [Medline: 16135720]
15. Vianello A, Fusello M, Gubian L, Rinaldo C, Dario C, Concas A, et al. Home telemonitoring for patients with acute exacerbation of chronic obstructive pulmonary disease: a randomized controlled trial. BMC Pulm Med 2016 Nov 22;16(1):157 [FREE Full text] [doi: 10.1186/s12890-016-0321-2] [Medline: 27876029]
16. Ho T, Huang C, Chiu H, Ruan S, Tsai Y, Yu C, HINT Study Group. Effectiveness of telemonitoring in patients with chronic obstructive pulmonary disease in Taiwan-a randomized controlled trial. Sci Rep 2016 Mar 31;6:23797 [FREE Full text] [doi: 10.1038/srep23797] [Medline: 27029815]
17. Gottlieb M, Marsaa K, Andreassen H, Strømstad G, Godtfredsen N. Feasibility of a telecare solution for patients admitted with COPD exacerbation: screening data from a pulmonary ward in a university hospital. Eur Clin Respir J 2014;1 [FREE Full text] [doi: 10.3402/ecrj.v1.24193] [Medline: 26557242]
18. Vitacca M, Rossin M, Assoni G, Baratti D, Zanardini M, Ruocco G, et al. Tele-assistance Respiratory card: feasibility of self-reporting in patients with severe COPD. Telemed J E Health 2013 Feb;19(2):99-103. [doi: 10.1089/tmj.2012.0089] [Medline: 23230820]
19. Almagro P, Barreiro B, Ochoa de Echaguen A, Quintana S, Rodríguez Carballeira M, Heredia JL, et al. Risk factors for hospital readmission in patients with chronic obstructive pulmonary disease. Respiration 2006;73(3):311-317. [doi: 10.1159/000088092] [Medline: 16155352]
20. Stone R, McMillan V, Mortier K, Holzhauer-Barrie J, Riordan J, Stone P. National COPD Audit Programme. Eur Respir J 2017;50(Suppl 61):PA4969. [doi: 10.1183/1393003.congress-2017.PA4969]
21. Farmer A, Williams V, Velardo C, Shah SA, Yu L, Rutter H, et al. Self-management support using a digital health system compared with usual care for chronic obstructive pulmonary disease: randomized controlled trial. J Med Internet Res 2017 May 03;19(5):e144 [FREE Full text] [doi: 10.2196/jmir.7116] [Medline: 28468749]
22. Hung MH, Chang MC, Kuo CW, Lin CM, Chuang LP, Kao KC. Four-weeks remote pulmonary rehabilitation protocol with mobile apps of real-time heart rate monitoring for gold category B/C/D: a study design. Respirology 2018 Nov 20;23:82-82. [doi: 10.1111/resp.13419_201]
23. Smaradottir B, Gerdes M, Fensli R, Martinez S. Usability evaluation of a COPD remote monitoring application. Stud Health Technol Inform 2015;210:845-849. [Medline: 25991274]
24. Cruz J, Brooks D, Marques A. Home telemonitoring in COPD: a systematic review of methodologies and patients' adherence. Int J Med Inform 2014 Apr;83(4):249-263. [doi: 10.1016/j.ijmedinf.2014.01.008] [Medline: 24529402]
25. Sanchez-Morillo D, Fernandez-Granero MA, Leon-Jimenez A. Use of predictive algorithms in-home monitoring of chronic obstructive pulmonary disease and asthma: A systematic review. Chron Respir Dis 2016 Aug;13(3):264-283 [FREE Full text] [doi: 10.1177/1479972316642365] [Medline: 27097638]
26. Chau JP, Lee DT, Yu DS, Chow AY, Yu W, Chair S, et al. A feasibility study to investigate the acceptability and potential effectiveness of a telecare service for older people with chronic obstructive pulmonary disease. Int J Med Inform 2012 Oct;81(10):674-682. [doi: 10.1016/j.ijmedinf.2012.06.003] [Medline: 22789911]
27. Koff PB, Jones RH, Cashman JM, Voelkel NF, Vandivier RW. Proactive integrated care improves quality of life in patients with COPD. Eur Respir J 2009 May;33(5):1031-1038 [FREE Full text] [doi: 10.1183/09031936.00063108] [Medline: 19129289]
28. Orme MW, Weedon AE, Saukko PM, Esliger DW, Morgan MD, Steiner MC, et al. Findings of the Chronic Obstructive Pulmonary Disease-Sitting and Exacerbations Trial (COPD-SEAT) in reducing sedentary time using wearable and mobile technologies with educational support: randomized controlled feasibility trial. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e84 [FREE Full text] [doi: 10.2196/mhealth.9398] [Medline: 29643055]
29. Houchen-Wolloff L, Orme M, Barradell A, Clinch L, Chaplin E, Gardiner N, et al. Web-Based Self-management Program (SPACE for COPD) for individuals hospitalized with an acute exacerbation of chronic obstructive pulmonary disease: nonrandomized feasibility trial of acceptability. JMIR Mhealth Uhealth 2021 Jun 11;9(6):e21728 [FREE Full text] [doi: 10.2196/21728] [Medline: 34114960]
30. Witry M, Comellas A, Simmering J, Polgreen P. The association between technology use and health status in a chronic obstructive pulmonary disease cohort: multi-method study. J Med Internet Res 2018 Apr 02;20(4):e125 [FREE Full text] [doi: 10.2196/jmir.9382] [Medline: 29610113]
31. Liu Y, Zhu SH, Wang GH, Ye F, Li PZ. Validity and reliability of multiparameter physiological measurements recorded by the Equivital LifeMonitor during activities of various intensities. J Occup Environ Hyg 2013;10(2):78-85. [doi: 10.1080/15459624.2012.747404] [Medline: 23259751]
32. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. Thorax 1999 Jul;54(7):581-586 [FREE Full text] [doi: 10.1136/thx.54.7.581] [Medline: 10377201]
33. Fan KG, Mandel J, Agnihotri P, Tai-Seale M. Remote patient monitoring technologies for predicting chronic obstructive pulmonary disease exacerbations: review and comparison. JMIR Mhealth Uhealth 2020 May 21;8(5):e16147 [FREE Full text] [doi: 10.2196/16147] [Medline: 32348262]

34. Rubio N, Parker RA, Drost EM, Pinnock H, Weir CJ, Hanley J, et al. Home monitoring of breathing rate in people with chronic obstructive pulmonary disease: observational study of feasibility, acceptability, and change after exacerbation. Int J Chron Obstruct Pulmon Dis 2017;12:1221-1231 [FREE Full text] [doi: 10.2147/COPD.S120706] [Medline: 28458534]
35. Downey C, Randell R, Brown J, Jayne DG. Continuous versus intermittent vital signs monitoring using a wearable, wireless patch in patients admitted to surgical wards: pilot cluster randomized controlled trial. J Med Internet Res 2018 Dec 11;20(12):e10802 [FREE Full text] [doi: 10.2196/10802] [Medline: 30538086]
36. Aliverti A. Wearable technology: role in respiratory health and disease. Breathe (Sheff) 2017 Jun;13(2):e27-e36 [FREE Full text] [doi: 10.1183/20734735.008417] [Medline: 28966692]
37. Soon S, Svavarsdottir H, Downey C, Jayne DG. Wearable devices for remote vital signs monitoring in the outpatient setting: an overview of the field. BMJ Innov 2020 Jan 14;6(2):55-71. [doi: 10.1136/bmjinnov-2019-000354]
38. Zheng Y, Ding X, Poon CCY, Lo BPL, Zhang H, Zhou X, et al. Unobtrusive sensing and wearable devices for health informatics. IEEE Trans Biomed Eng 2014 May;61(5):1538-1554 [FREE Full text] [doi: 10.1109/TBME.2014.2309951] [Medline: 24759283]
39. Leenen JPL, Leerentveld C, van Dijk JD, van Westreenen HL, Schoonhoven L, Patijn GA. Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: systematic review. J Med Internet Res 2020 Jun 17;22(6):e18636 [FREE Full text] [doi: 10.2196/18636] [Medline: 32469323]
40. Wedzicha JA, Seemungal TA. COPD exacerbations: defining their cause and prevention. Lancet 2007 Sep 01;370(9589):786-796 [FREE Full text] [doi: 10.1016/S0140-6736(07)61382-8] [Medline: 17765528]
41. Cuddy J, Ruby B, Santee W, Karis A. Hidalgo Equivital Physiological Monitor Product Review and Data Summary. U.S. Army Research Institute of Environmental Medicine. 2008. URL: https://apps.dtic.mil/sti/pdfs/ADA491504.pdf [accessed 2022-01-17]
42. Phillips M, Beach J, Cathey R, Lockert J, Satterfield W. Reliability and validity of the Hexoskin Wearable Body Metrics Telemetry Shirt. J Sport Hum Perform 2017;5(2):8. [doi: 10.12922/jshp.v5i2.108]
43. Kim J, Roberge R, Powell J, Shafer A, Jon Williams W. Measurement accuracy of heart rate and respiratory rate during graded exercise and sustained exercise in the heat using the Zephyr BioHarness. Int J Sports Med 2013 Jun;34(6):497-501 [FREE Full text] [doi: 10.1055/s-0032-1327661] [Medline: 23175181]
44. McAuley HJC, Harvey-Dunstan TC, Craner M, Richardson M, Singh SJ, Steiner MC, et al. Longitudinal changes to quadriceps thickness demonstrate acute sarcopenia following admission to hospital for an exacerbation of chronic respiratory disease. Thorax 2021 Jul;76(7):726-728. [doi: 10.1136/thoraxjnl-2020-215949] [Medline: 33443216]
45. Vilaró J, Ramirez-Sarmiento A, Martínez-Llorens JMA, Mendoza T, Alvarez M, Sánchez-Cayado N, et al. Global muscle dysfunction as a risk factor of readmission to hospital due to COPD exacerbations. Respir Med 2010 Dec;104(12):1896-1902 [FREE Full text] [doi: 10.1016/j.rmed.2010.05.001] [Medline: 20541383]
46. Spruit MA, Gosselink R, Troosters T, Kasran A, Gayan-Ramirez G, Bogaerts P, et al. Muscle force during an acute exacerbation in hospitalised patients with COPD and its relationship with CXCL8 and IGF-I. Thorax 2003 Sep;58(9):752-756 [FREE Full text] [doi: 10.1136/thorax.58.9.752] [Medline: 12947130]
47. Abnormal Heart Rate Readings During Exercise. Polar. URL: https://support.polar.com/en/support/FAQs/Abnormal_Heart_Rate_Readings_During_Exercise?category= [accessed 2022-01-17]

Abbreviations

AECOPD: acute exacerbation of chronic obstructive pulmonary disease
COPD: chronic obstructive pulmonary disease
HR: heart rate
NHS: National Health Service
NIHR: National Institute for Health Research
PA: physical activity
PPI: patient and public involvement
PR: pulmonary rehabilitation
RR: respiratory rate
SEM: sensor electronics module
SpO_2: oxygen saturation
ST: skin temperature
VSM: vital signs monitoring

Edited by A Kushniruk; submitted 30.04.21; peer-reviewed by F Lai, JA Krishnan; comments to author 04.07.21; revised version received 30.07.21; accepted 26.08.21; published 16.02.22.

Please cite as:
Hawthorne G, Greening N, Esliger D, Briggs-Price S, Richardson M, Chaplin E, Clinch L, Steiner MC, Singh SJ, Orme MW
Usability of Wearable Multiparameter Technology to Continuously Monitor Free-Living Vital Signs in People Living With Chronic Obstructive Pulmonary Disease: Prospective Observational Study
JMIR Hum Factors 2022;9(1):e30091
URL: https://humanfactors.jmir.org/2022/1/e30091
doi:10.2196/30091
PMID:35171101

Original Paper

User Perceptions of Different Vital Signs Monitor Modalities During High-Fidelity Simulation: Semiquantitative Analysis

Samira Akbas[1], MD; Sadiq Said[1], MD; Tadzio Raoul Roche[1], MD; Christoph B Nöthiger[1], MD; Donat R Spahn[1], MD; David W Tscholl[1], MD; Lisa Bergauer[1], MD

Institute of Anesthesiology, University of Zurich, University Hospital Zurich, Zurich, Switzerland

Corresponding Author:
David W Tscholl, MD
Institute of Anesthesiology
University of Zurich
University Hospital Zurich
Raemistrasse 100
Zurich, 8091
Switzerland
Phone: 41 432539451
Email: david.tscholl@usz.ch

Abstract

Background: Patient safety during anesthesia is crucially dependent on the monitoring of vital signs. However, the values obtained must also be perceived and correctly classified by the attending care providers. To facilitate these processes, we developed Visual-Patient-avatar, an animated virtual model of the monitored patient, which innovatively presents numerical and waveform data following user-centered design principles. After a high-fidelity simulation study, we analyzed the participants' perceptions of 3 different monitor modalities, including this newly introduced technique.

Objective: The aim of this study was to collect and evaluate participants' opinions and experiences regarding 3 different monitor modalities, which are Visual-Patient-avatar, Split Screen (avatar and Conventional monitor alongside each other), and Conventional monitor after using them during simulated critical anesthetic events.

Methods: This study was a researcher-initiated, single-center, semiquantitative study. We asked 92 care providers right after finishing 3 simulated emergency scenarios about their positive and negative opinions concerning the different monitor modalities. We processed the field notes obtained and derived the main categories and corresponding subthemes following qualitative research methods.

Results: We gained a total of 307 statements. Through a context-based analysis, we identified the 3 main categories of "Visual-Patient-avatar," "Split Screen," and "Conventional monitor" and divided them into 11 positive and negative subthemes. We achieved substantial interrater reliability in assigning the statements to 1 of the topics. Most of the statements concerned the design and usability features of the avatar or the Split Screen mode.

Conclusions: This study semiquantitatively reviewed the clinical applicability of the Visual-Patient-avatar technique in a high-fidelity simulation study and revealed the strengths and limitations of the avatar only and Split Screen modality. In addition to valuable suggestions for improving the design, the requirement for training prior to clinical implementation was emphasized. The responses to the Split Screen suggest that this symbiotic modality generates better situation awareness in combination with numerical data and accurate curves. As a subsequent development step, a real-life introduction study is planned, where we will test the avatar in Split Screen mode under actual clinical conditions.

(JMIR Hum Factors 2022;9(1):e34677) doi:10.2196/34677

KEYWORDS

avatar; patient monitoring; semiquantitative research; simulation study; situation awareness; user-centered design; visual-patient-avatar

Introduction

Although perioperative mortality directly attributable to anesthesia is low in high-income countries and has significantly declined over the last 50 years, the World Health Organization describes anesthesiologic and surgical complications as the leading cause of preventable perioperative morbidity and mortality [1-4].

Among all anesthesia complications leading to permanent brain damage or death, two-thirds are caused by inaccurate situation awareness. This concept developed by Mica Endsley comprises a chain of information processing including the three core levels of perception (level 1), comprehension (level 2), and projection (level 3), whereby level 1 is most frequently affected [5-7]. For appropriate decision-making and thus avoiding errors, a situation must be recognized, its severity assessed, and the correct next steps taken while anticipating future progress. This cognitive process is influenced by individual factors such as experience and environmental resources. Well-established methods such as perioperative checklists were developed with the intention of improving environmental resources [8]. In addition, new tools are needed to impact situation awareness positively and thus reduce perioperative anesthesiologic complications.

Hence, we developed Visual-Patient-avatar as a beneficial environmental factor on situation awareness in patient monitoring. This avatar-based visualization on a patient monitor displays an animated model of the measured numerical parameters combining principles of logic and user-centered design [9]. Previous computer-based studies have shown that more vital signs were observed when using this new technique, subjective diagnostic confidence increased, and perceived workload declined compared to conventional patient monitoring [10-12]. However, the use of Visual-Patient-avatar in a high-fidelity simulation study has not yet been analyzed, including its qualitative aspects.

This study aims to collect and assess the opinions and experiences of participants concerning the three different patient monitoring modalities, which are (1) Visual-Patient-avatar, (2) Split Screen (avatar and conventional patient monitoring side by side), and (3) conventional, after using them in simulated critical anesthesia events [13]. We sought to capture the advantages and disadvantages of the different monitor settings to foster the avatar's development and, in the future, facilitate its implementation in everyday clinical practice.

Methods

Approval and Consent

The Cantonal Ethics Committee of Zurich in Switzerland issued a declaration of no objection after reviewing the study protocol (Business Management System for Ethics Committees Req-2020-00059). All participants signed written informed consent for the use of their data for research purposes and participated voluntarily without any financial compensation.

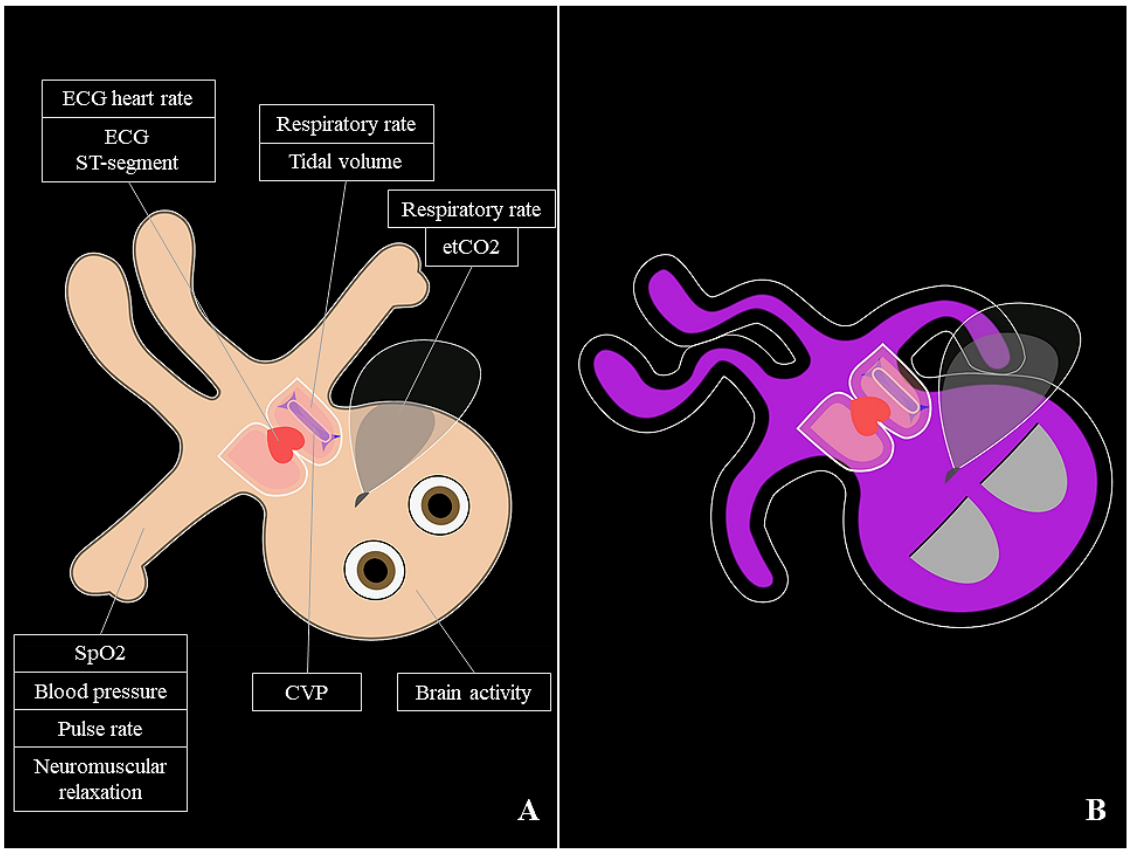
Study Design

This is a researcher-initiated, single-center, semiquantitative study investigating physicians' and nurses' perceptions of using Visual-Patient-avatar in simulated critical anesthesia events. We conducted this study at the University Hospital of Zurich in Switzerland, in May 2020. We included the same 104 care providers grouped in 52 teams of a recently published study that evaluated avatar-based patient monitoring in a high-fidelity simulation study [13].

Previous Avatar-Based Patient Monitoring Simulation Study and Participant Interviews

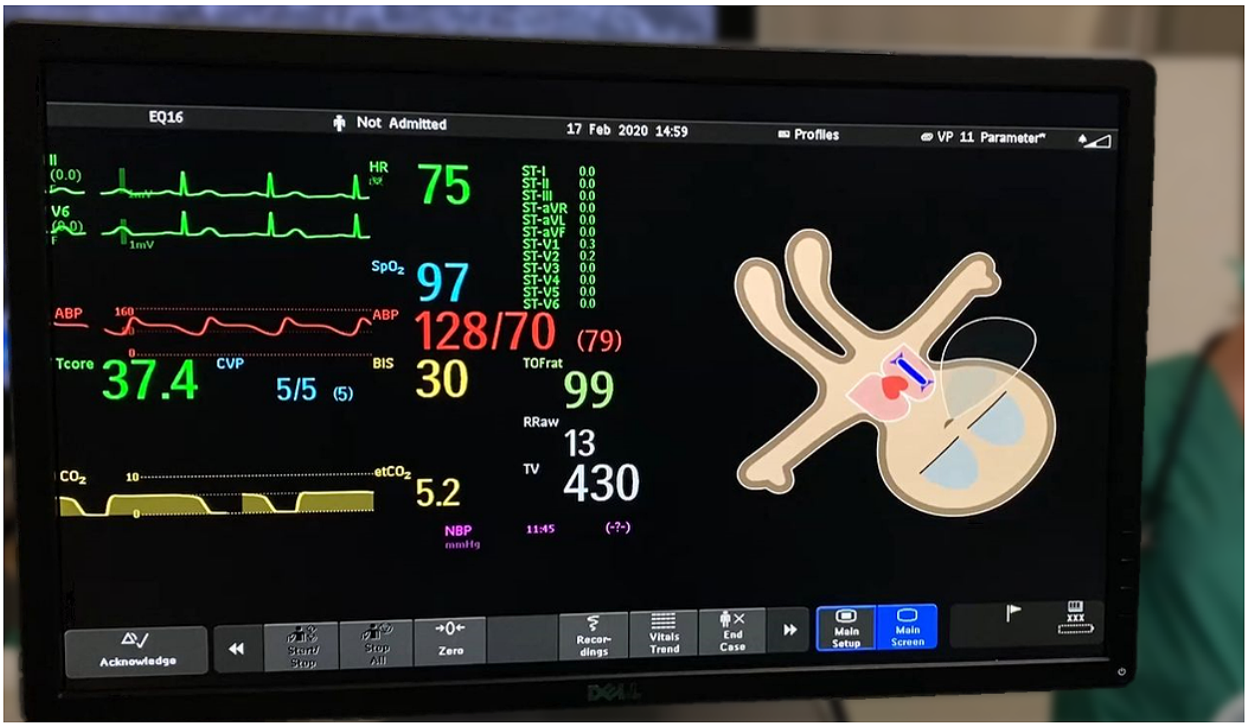
This recently published study showed noninferiority of Split Screen compared with Conventional monitoring for performance during anesthesia crisis events. The probability of communicating the correct reason for the emergency was increased using the Visual-Patient-avatar as the monitor modality [13]. Figure 1 shows the 13 available vital signs and an example with possible deviations and additionally. Part (A) depicts an awake patient with vital signs within normal range. The avatar's body pulsates during patient monitoring, whereby the frequency and extension indicate the pulse rate and blood pressure, respectively. In part (B), we demonstrate a desaturated (purple color), deeply sedated (eyes closed) patient with muscle relaxation (floppy extremities). Hypotension is represented by the gap between the purple body and the white boundary line. If the body temperature leaves the normal range, ice crystals or heat waves become visible around the avatar. Additionally, Multimedia Appendix 1 provides an animated version of the 2 examples.

Figure 1. Two examples of Visual-Patient-avatar used during the high-fidelity simulation. CVP: central venous pressure; ECG: electrocardiogram; etCO2: end-tidal carbon dioxide; SpO2: peripheral oxygen saturation; ST: ST-segment.



After a short briefing and a training scenario, the participants completed 3 different emergency scenarios, each with 1 of the following 3 different monitor modalities: only Visual-Patient-avatar, Split Screen, and only conventional, number-based and waveform-based monitor. Figure 2 illustrates an example of a Split Screen display during simulation, and the video in Multimedia Appendix 2 shows a recording of a simulation scenario. After completing all scenarios, we asked the following 2 questions: “What do you like about the monitor settings? Eg, particular strengths?” and “What do you dislike about the monitor settings? Eg, potential problems, limitations?” The study authors TRR and SS recorded the participants’ responses as field notes on an iPad (Apple Inc). The participants reviewed the final field note transcripts, modifying or adding to them if warranted.

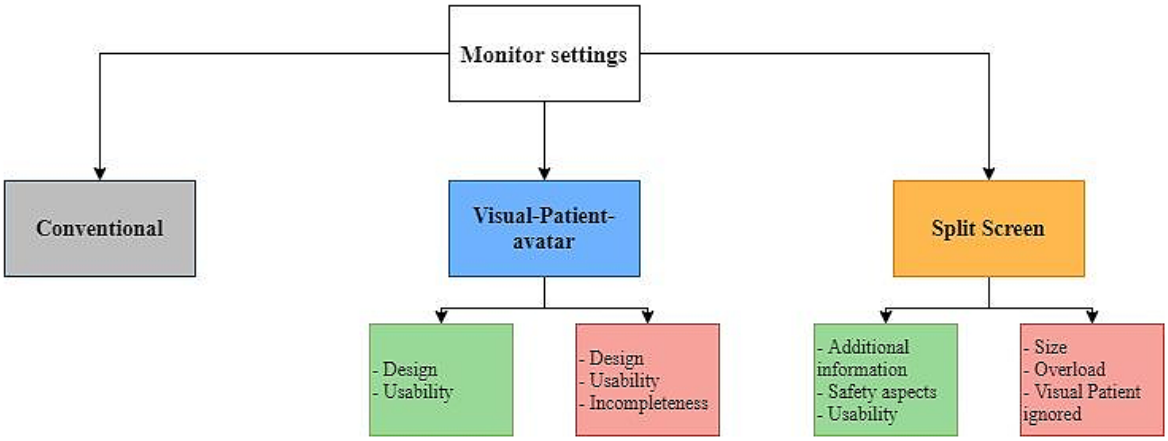
Figure 2. Example of a Split Screen display during simulation with the Conventional monitor on the left and the Visual-Patient-avatar on the right side. The beige skin tone corresponds to a normal peripheral oxygen saturation, and closed eyes imply a sedated patient.



Semiquantitative Analysis

After collecting all answers, we translated them from German to English using an online translating service, Deepl (DeepL GmbH). In Multimedia Appendix 3, we provide the complete translated field notes. There were no comments made from 12 participants. To gain a first impression by identifying frequently mentioned terms, we excluded filler words such as “and” or “the” and performed a word count using Microsoft Word (Microsoft Corporation). Although the word count does not provide information about the content of individual statements, this approach helped us to identify similar expressions. Subsequently, we grouped the statements using the template approach, identified main topics, and generated a coding tree, which we modified until all essential and frequent statements could be classified [14,15]. According to the recommendations of reporting qualitative research, study authors SA and LB, who were not involved in the interview process, evaluated the statements independently of each other, using the final coding tree displayed in Figure 3, which was created using draw.io (Seibert Media GmbH) [16-18]. Before determining a joint code in case of disagreement, we calculated the interrater reliability to validate the rating.

Figure 3. Hierarchical coding tree concerning user perceptions of the 3 different monitor modalities. The green boxes include positive subthemes of the respective major topic. The red boxes contain negative subthemes of the corresponding major topic.



Statistical Analysis

We report the number of statements and their percentages relating to the superior topics. To manage our data and to generate the figures, we used Microsoft Word and Microsoft Excel. To quantify the interrater reliability, when assigning the individual statements to a particular topic of the final coding tree, we calculated the Cohen kappa using R, version 4.0.5 (R Foundation for Statistical Computing) [19]. To estimate percent agreement, we used Microsoft Excel.

Results

Participant and Field Notes Characteristics

We acquired field notes between May 4, 2020, and May 28, 2020. All participants were employees of the University Hospital of Zurich. Of the total 104 participants, 92 (88%) senior physicians, resident physicians, nurse anesthetists, and trainee nurse anesthetists took part in the interview process right after completing the simulation scenarios. Table 1 provides a detailed description of the study and participant characteristics.

Analyzing the field notes obtained, we identified 329 individual statements consisting of 2807 words. Of 329 statements, 22 (7%) were not comprehensible to us in terms of content even after several discussions, so we classified them as "not codable." Statements in this category included subjective opinions such as "I like it" (participant #13.1). The remaining 307 statements were analyzed semiquantitatively, which allows the calculation of the proportions of individual statements among the main topics and subthemes without applying statistical tests [20,21]. Overall, the ratio of statements to the question, "What do you like about the monitor settings" (144/307, 47%) compared to the negative perceptions (163/307, 53%) was balanced.

Table 1. Study and participant characteristics in detail (n=92).

Study and participant characteristics	Values
Participants who submitted field notes, n (%)	92 (88)
Female participants, n (%)	46 (50)
Senior physicians, n (%)	14 (15)
Resident physicians, n (%)	33 (36)
Nurse anesthetists, n (%)	30 (33)
Trainee nurse anesthetists, n (%)	17 (16)
Total anesthesia experience (years), mean (IQR)	6.6 (1.5-8)

Semiquantitative Analysis

Beginning our semiquantitative analysis, we performed word counts to expose potential main themes. The analysis revealed that "Visual Patient" was the most frequently occurring term in the field notes obtained to answer both the positive (35 times in 144 statements) and negative questions (38 times in 163 statements).

Based on using qualitative research methods and testing 3 monitor modalities, the final coding tree contains the 3 main categories of Visual-Patient-avatar, Split Screen, and Conventional monitor, as well as 4 main topics with 11 subthemes. When independently assigning all 327 statements received to 1 of these topics, the study authors SA and LB achieved 80% interrater agreement with a substantial Cohen kappa of 0.78 [22]. In the case of differently coded statements, a review and joint assignment followed to achieve 100% interrater agreement after the second round of coding. Figure 4 visualizes the percentage distribution of all statements among the different categories. The 3 main categories are located in the innermost circle. The associated major topics and subthemes are displayed hierarchically toward the outside. Table 2 outlines the major topics with examples. In the subsequent sections, we describe the individual categories in detail with percentages and examples. The calculations refer to the codable statements.

Figure 4. Sunburst diagram to reflect the user perceptions of the 3 different monitor settings. The width of a section represents the respective percentage of the topic on all given statements (N=307). Ignored VP: ignored Visual-Patient-avatar.

Table 2. The major topics with participant count, percentages, and examples.

Major topics and subthemes	Examples
Visual-Patient-avatar positive (61/307, 20%)	
Design	• Especially oxygenation and body temperature well displayed. [#[a]11.2] • Integration of all values on one avatar. [#47.2]
Usability	• Information simplified by Visual-Patient-avatar. [#13.2] • Overview of relevant parameters at a glance through Visual-Patient-avatar. [#25.2]
Visual-Patient-avatar negative (99/307, 32%)	
Design	• The blood pressure feature was not easy to understand for me. [#24.2] • Visual-Patient-avatar: head too large in contrast to heart and lung. [#25.1]
Usability	• Visual-Patient-avatar takes some time getting used to; not entirely intuitive at first. [#29.2] • At the moment still difficult but with potential. [#38.1]
Incompleteness	• Numbers and ranks are missing. [#10.1] • Lacking information quantification with the Visual-Patient-avatar. [#48.1]
Split Screen positive (79/307, 26%)	
Additional information	• Additional information by Visual-Patient-avatar. [#14.2] • I like the combination of new and old monitoring. [#18.2]
Safety aspects	• With Visual-Patient-avatar changes faster visible than with numbers or curves. [#13.1] • More safety. [#41.1]
Usability	• Split monitoring helps to focus. [#10.1] • I prefer the split monitor, Visual-Patient-avatar as first initial diagnosis—quantification via Conventional monitoring. [#16.1]
Split Screen negative (60/307, 20%)	
Size	• Info partly displayed a bit small. [#13.2] • Needs appropriate monitor size. [#24.1]
Overload	• Too much information at once in the emergency situation. [#12.2] • Screen very full. [#30.2]
Visual-Patient-avatar ignored	• Looked at numbers. [#9.1] • I barely looked at the Visual-Patient-avatar. [#18.2]
Conventional monitor (8/307, 3%)	• I want to see the details or parameter more precisely. I prefer the “usual” monitor view. [#26.1]

[a]Participant number.

Statements About Visual-Patient-Avatar

We assigned 160 of 307 (52%) statements to the main category Visual-Patient-avatar. Through inductive free coding, the 2 major topics, Visual-Patient-avatar positive (61/307, 20%) and Visual-Patient-avatar negative (99/307, 32%), were revealed.

We divided the positive major topic into the 2 subthemes, design (26/307, 8%) and usability (32/307, 10%). Concerning design features, the participants distinguished the simplified (participant #13.2) and realistic (participant #17.1) appearance of the avatar. Participant #11.2 outlined that “Especially oxygenation and body temperature is illustrated well.” The participants also recognized advantages in terms of usability. They found that Visual-Patient-avatar is “Intuitively understandable” (participant #31.1), “Gives a good overview” (participant #34.1), and helps to grasp the situation quickly (participant #37.1). We allocated more common statements such as “Integration of all values on one avatar” (participant #47.2) to the major topic Visual-Patient-avatar positive (3/307, 1%).

Regarding negative properties of the avatar, the participants’ responses depicted design (15/307, 5%), usability (45/307, 15%), and incompleteness (39/307, 13%) as subthemes. For participant #43.2, the thorax displayed too small, and the vena cava representation was unclear. Others raised concerns about possible misinterpretations (participant #46.1) because of the unfamiliar (participant #48.2) and confusing (participant #48.1) vital sign presentation within Visual-Patient-avatar technique. Without concrete values (participant #35.2) and curves such as the electrocardiogram (participant #42.2), the avatar did not help in solving the emergency scenarios.

Statements About the Split Screen

In 139 of 307 (46%) statements, the participants noticed this main category, which we classified into the major topics Split Screen positive (79/307, 26%) and Split Screen negative (60/307, 20%).

In interrater consent, the positive major topic included the 3 subthemes additional information (29/307, 9%), safety aspects (21/307, 7%), and usability (11/307, 4%). Through "Increasing attention" (participant #10.1), "Faster recognition of changes" (participant #29.2), and the "Quick overview" (participant #45.2), the participants perceived a higher level of safety. Several participants found the Split Screen mode overall "Helpful" (participant #47.1) and "Effective" (participant #17.1) in its use. We allocated responses that generally considered the combination advantageous to the major topic, positive Split Screen (18/307, 6%).

The negative major topic concerning Split Screen enclosed the 3 subthemes, "size," "overload," and "Visual-Patient-avatar ignored." These were named, respectively, in 23/307 (7%), 23/307 (7%), and 9/307 (3%) statements. The participants criticized the small display and thus the difficulty of detecting details of the curves (participant #20.1) and the Visual-Patient-avatar (participant #21.2). Furthermore, they claimed the Split Screen to be crowded (participant #30.2), and that there is "Too much information at once in the emergency" (participant #12.2). In addition, the analysis of the field notes discovered that several participants ignored the avatar. General annotations such as "Not sure about the added benefit" (participant #48.1) were assigned to the major topic, negative Split Screen (5/307, 2%).

Statements About the Conventional Monitor

A small number of the field notes referred to the main category Conventional monitor (8/307, 3%). Some participants just stated that they "Prefer the usual monitor view" (participant #26.1). Furthermore, we grouped responses that mentioned the familiar audio support in this main category.

Discussion

Principal Findings

This semiquantitative single-center study explored the impressions of anesthesia personnel when using the existing Conventional monitor compared with the new modality Visual-Patient-avatar—either the avatar only or the Split Screen variant. User perceptions can uncover improvement opportunities, and their consideration is essential for the success of new medical techniques. We assigned most of the statements to the main category Visual-Patient-avatar, highlighting positive characteristics and negative features such as the absence of quantitative data. Many annotations also evaluated the Split Screen modality, while only a few participants commented on the well-known Conventional monitor. The latter seems coherent as Visual-Patient-avatar is a novelty and thus attention catching.

The avatar's development was guided by the idea of providing a monitor tool that improves situation awareness through its user-centered design principles. Following the definition of a user-centered design through Mica Endsley [9], many participants stated the avatar technique to include beneficial design and usability features such as being simplified and intuitive (positive Visual-Patient-avatar: 61/307, 20%). When used with the Conventional monitor in Split Screen mode, Visual-Patient-avatar increases attention and provides a quick overview. Possible changes in vital signs can then be quantified using the conventional display (safety aspects: 21/307, 7%). The aspect of time saving through faster detection is essential in patient care, as for example, postoperative renal dysfunction is related to the overall duration of hypotension during general anesthesia [23].

Many participants claimed the missing numbers and curves when using only the avatar makes a more precise diagnosis impossible (incompleteness: 39/307, 13%). In Visual-Patient-avatar, the data for each vital sign is preprocessed to show different states (no data, too low, normal, or too high), aiming to reduce complexity. We understand the technology as a supplement, which cannot replace the Conventional monitor; however, it can improve care providers' situation awareness by presenting information that is easy to perceive and comprehend.

On the question, "What did you dislike about the monitor settings?", this analysis found that size (23/307, 7%) and overload (23/307, 7%) were the main critical points concerning the Split Screen mode. During the simulation study, the scenarios run on 12-inch patient monitors (Philips IntelliVue MX500; Koninklijke Philips NV, Amsterdam, The Netherlands). However, the technique for the real-life clinical implementation is compatible with the Philips IntelliVue MX 550 monitor, which offers a larger display of 15 inches. This fact can mitigate the criticism, but it is known that a high information load can have a detrimental effect on the ability to set priorities and can confuse the individual [24]. This would contradict the basic idea of Visual-Patient-avatar and must be kept in mind.

The impression of an overload could also occur because the technique of the Visual-Patient-avatar and its implementation as Split Screen variant is new and therefore cognitively demanding. Accordingly, several participants mentioned being unfamiliar with the avatar, whether used individually or in Split Screen mode. Upon introduction of the new technique into clinical routine, all users will receive education and training lessons. Nevertheless, it will take time to get used to the new monitor modalities and fully implement them mentally, as especially very experienced care providers have been working with the Conventional monitor modality for decades. The successful implementation of new techniques can be demonstrated by the sonographically guided insertion of central venous catheters [25]. After initial skepticism, this method is nowadays preferred both in the literature and clinically, as the complication rate is lower than with landmark-guided puncture [26].

To achieve a high level of user acceptance, an intuitive interface and cognitive ease are crucial points [27]. Visual-Patient-avatar presents the information close to clinical reality. For example, the avatar's skin turns purple in case of hypoxemia, or its eyes are open when the brain-activity sensor detects a high signal. The participants' appreciation of the realistic and clear vital

sign display (design: 26/307, 8%) is in line with the results of the study by Wachter and colleagues [28], which shows that an anatomically related interface is particularly intuitive. However, together with design and technical specialists from Philips (Koninklijke Philips NV), an intensive redesign process was carried out to improve weaknesses in the design, such as the vena cava display. Even though some steps are still needed until clinical introduction, we expect visualization techniques to have a great future in medicine after this study. It is encouraging that Hamilton Medical AG provides the "dynamic lung" in ventilators to visualize specific lung parameters [29].

Strengths and Limitations

This study has several limitations. In qualitative and semiquantitative analysis, the structure and results are developed inductively and cannot be applied to a broader population as it does not investigate statistical significance. Nevertheless, this approach allowed us to gain firsthand perceptions and experiences from our participants right after using the different monitor modalities. Generally, a nonquantitative assessment stays close to the participants' point of view, implying a certain subjectivity [30,31]. However, this is put into perspective by the high number of participants and their diversity. As a single-center study, possible selection bias cannot be excluded. It is conceivable that the results vary under different circumstances.

To date, only computer-based studies have been conducted with the Visual-Patient-avatar [10,11]. One of the strengths is the high-fidelity simulation, which made it possible to test the new technique realistically during anesthesiologic emergency scenarios and to derive conclusions for its use under clinical conditions [32]. Thus, we obtained the first opinions on Visual-Patient-avatar directly after experiencing the urgency of emergent patient treatment. These findings greatly impact the further development of the technique up to the point of clinical implementation.

Conclusion

We designed this study to determine care providers' perceptions concerning monitor modalities incorporating the Visual-Patient-avatar technique. One of the key findings was that the participants experienced the avatar technique's underlying design principles and characteristics positively under active use in the context of a high-fidelity simulation. This insight complements those of earlier studies using the Visual-Patient-avatar technique in computer-based studies [33]. The participants confirmed the value of the Split Screen mode through its combination of visual impressions and simultaneous quantification with numerical parameters. This monitor variant, planned for future clinical implementation, gives a quick overview and draws attention to changes specified by the conventional part. The next step in the development is a planned real-life introduction study of the avatar in Split Screen mode under actual clinical conditions. This modality's weakness, based on the large amount of information displayed, will be reflected in the further planning process and will be reviewed through future studies. By testing the avatar in a simulated clinical environment for the first time, we are taking a significant step toward our vision: to help care providers in situations of high cognitive load to better prioritize information and thus positively influence decision-making for the patient's benefit.

Acknowledgments

The authors are thankful to the study participants for their time and effort. Furthermore, the authors thank biostatistician Julia Braun-Gruebel for her statistical advice regarding this manuscript. This study was funded by the Institute of Anesthesiology, University Hospital of Zurich, Switzerland.

Availability of Data and Materials

The data sets used and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conception and design: SA, SS, TRR, CBN, DRS, DWT, and LB. Acquisition of data: SA, SS, TRR, CBN, DWT, and LB. Analysis and interpretation of data: SA, DRS, DWT, and LB. Drafting of the article or critical revision of the manuscript: SA, SS, TRR, CBN, DRS, DWT, and LB. Final approval of the version to be published: SA, SS, TRR, CBN, DRS, DWT, and LB.

Conflicts of Interest

DWT and CBN are designated inventors of Visual-Patient-avatar technology, for which the University of Zurich holds various patents and trademarks. There are cooperation and licensing agreements with Philips Medizin Systeme Böblingen GmbH, Böblingen, Germany; Koninklijke Philips NV, Amsterdam, The Netherlands; Philips Research/Philips Electronics Nederland BV, Eindhoven, The Netherlands, Philips North America, Cambridge, MA, USA. Under these agreements, DWT receives grant support, and DWT and CBN receive travel support and may receive royalties. DWT, CBN, and DRS are designated inventors of Visual Clot technology, for which the University of Zurich holds various patents and trademarks. The University of Zurich signed a letter of intent for a cooperation and licensing agreement with Instrumentation Laboratory Company/Werfen Corporation, Bedford, MA, USA and Barcelona, Spain. Under resulting agreements, they may receive royalties. DWT, CBN, and DRS received travel support for consulting Instrumentation Laboratory, Bedford, MA, USA. DWT received honoraria and travel support from the Swiss Foundation for Anesthesia Research, Zurich, Switzerland. DRS's academic department is receiving grant support from the Swiss National Science Foundation, Berne, Switzerland, the Swiss Society of Anesthesiology and Perioperative Medicine,

Berne, Switzerland, the Swiss Foundation for Anesthesia Research, Zurich, Switzerland, Vifor SA, Villars-sur-Glâne, Switzerland, and Vifor (International) AG, St. Gallen, Switzerland. DRS is cochair of the ABC-Trauma Faculty, sponsored by unrestricted educational grants from Novo Nordisk Health Care AG, Zurich, Switzerland, CSL Behring GmbH, Marburg, Germany, LFB Biomédicaments, Courtaboeuf Cedex, France, and Octapharma AG, Lachen, Switzerland. DRS received honoraria or travel support for consulting or lecturing from the following: Danube University of Krems, Austria; US Department of Defense, Washington; European Society of Anesthesiology and Intensive Care, Brussels, BE; Korean Society for Patient Blood Management, Seoul, Korea; Korean Society of Anesthesiologists, Seoul, Korea; Network for the Advancement of Patient Blood Management, Haemostasis and Thrombosis, Paris, France; Alexion Pharmaceuticals Inc, Boston, MA, USA; Bayer AG, Zürich, Switzerland; B. Braun Melsungen AG, Melsungen, Germany; CSL Behring GmbH, Hattersheim am Main, Germany and Berne, Switzerland; Celgene International II Sàrl, Couvet, Switzerland; Daiichi Sankyo AG, Thalwil, Switzerland; Haemonetics, Braintree, MA, USA; Instrumentation Laboratory (Werfen), Bedford, MA, USA; LFB Biomédicaments, Courtaboeuf Cedex, France; Merck Sharp & Dohme, Kenilworth, New Jersey, USA; Novo Nordisk Health Care AG, Zurich, Switzerland; PAION Deutschland GmbH, Aachen, Germany; Pharmacosmos A/S, Holbaek, Denmark, Pfizer AG, Zürich, Switzerland; Pierre Fabre Pharma, Alschwil, Switzerland; Portola Schweiz GmbH, Aarau, Switzerland, Roche Diagnostics International Ltd, Reinach, Switzerland; Sarstedt AG & Co., Sevelen, Switzerland and Nümbrecht, Germany; Shire Switzerland GmbH, Zug, Switzerland; Tem International GmbH, Munich, Germany; Vifor Pharma, Munich, Germany; Neuilly sur Seine, France and Villars-sur-Glâne, Switzerland, Vifor (International) AG, St. Gallen, Switzerland; and Zuellig Pharma Holdings, Singapore, Singapore.

Multimedia Appendix 1
Video with two examples of an animated Visual-Patient-avatar.
[MP4 File (MP4 Video), 27521 KB - humanfactors_v9i1e34677_app1.mp4]

Multimedia Appendix 2
Video of simulations scenario presenting Visual-Patient-avatar and Conventional monitor.
[MP4 File (MP4 Video), 60844 KB - humanfactors_v9i1e34677_app2.mp4]

Multimedia Appendix 3
The translated field notes of 92 participants. The brackets [] indicate which parts were each assessed as one statement.
[DOCX File , 38 KB - humanfactors_v9i1e34677_app3.docx]

References

1. WHO Guidelines for Safe Surgery 2009. World Health Organization. 2009. URL: http://apps.who.int/iris/bitstream/handle/10665/44185/9789241598552_eng.pdf;jsessionid=BC88C8E245C68B57989D3EBAFC3FC3EE?sequence=1 [accessed 2022-03-07]
2. Irwin MG, Kong VKF. Quantifying and communicating peri-operative risk. Anaesthesia 2014 Dec;69(12):1299-1303 [FREE Full text] [doi: 10.1111/anae.12881] [Medline: 25271597]
3. Li G, Warner M, Lang BH, Huang L, Sun LS. Epidemiology of anesthesia-related mortality in the United States, 1999-2005. Anesthesiology 2009 Apr;110(4):759-765 [FREE Full text] [doi: 10.1097/aln.0b013e31819b5bdc] [Medline: 19322941]
4. Bainbridge D, Martin J, Arango M, Cheng D, Evidence-based Peri-operative Clinical Outcomes Research (EPiCOR) Group. Perioperative and anaesthetic-related mortality in developed and developing countries: a systematic review and meta-analysis. Lancet 2012 Sep 22;380(9847):1075-1081. [doi: 10.1016/S0140-6736(12)60990-8] [Medline: 22998717]
5. Schulz CM, Krautheim V, Hackemann A, Kreuzer M, Kochs EF, Wagner KJ. Situation awareness errors in anesthesia and critical care in 200 cases of a critical incident reporting system. BMC Anesthesiol 2016 Jan 16;16:4 [FREE Full text] [doi: 10.1186/s12871-016-0172-7] [Medline: 26772179]
6. Schulz CM, Burden A, Posner KL, Mincer SL, Steadman R, Wagner KJ, et al. Frequency and Type of Situational Awareness Errors Contributing to Death and Brain Damage: A Closed Claims Analysis. Anesthesiology 2017 Aug;127(2):326-337 [FREE Full text] [doi: 10.1097/ALN.0000000000001661] [Medline: 28459735]
7. Schulz CM, Endsley MR, Kochs EF, Gelb AW, Wagner KJ. Situation awareness in anesthesia: concept and research. Anesthesiology 2013 Mar;118(3):729-742 [FREE Full text] [doi: 10.1097/ALN.0b013e318280a40f] [Medline: 23291626]
8. WHO's patient-safety checklist for surgery. The Lancet 2008 Jul 05;372(9632):1. [doi: 10.1016/S0140-6736(08)60964-2] [Medline: 18603137]
9. Endsley M. Designing for situation awareness: an approach to user-centered design. Boca Raton, Florida, US: CRC Press; Jan 18, 2012:1-396.
10. Garot O, Rössler J, Pfarr J, Ganter MT, Spahn DR, Nöthiger CB, et al. Avatar-based versus conventional vital sign display in a central monitor for monitoring multiple patients: a multicenter computer-based laboratory study. BMC Med Inform Decis Mak 2020 Feb 10;20(1):26 [FREE Full text] [doi: 10.1186/s12911-020-1032-4] [Medline: 32041584]

11. Tscholl DW, Handschin L, Neubauer P, Weiss M, Seifert B, Spahn DR, et al. Using an animated patient avatar to improve perception of vital sign information by anaesthesia professionals. Br J Anaesth 2018 Sep;121(3):662-671 [FREE Full text] [doi: 10.1016/j.bja.2018.04.024] [Medline: 30115265]
12. Tscholl DW, Rössler J, Handschin L, Seifert B, Spahn DR, Nöthiger CB. The Mechanisms Responsible for Improved Information Transfer in Avatar-Based Patient Monitoring: Multicenter Comparative Eye-Tracking Study. J Med Internet Res 2020 Mar 16;22(3):e15070 [FREE Full text] [doi: 10.2196/15070] [Medline: 32175913]
13. Roche TR, Said S, Braun J, Maas EJC, Machado C, Grande B, et al. Avatar-based patient monitoring in critical anaesthesia events: a randomised high-fidelity simulation study. Br J Anaesth 2021 May;126(5):1046-1054. [doi: 10.1016/j.bja.2021.01.015] [Medline: 33879327]
14. Brooks J, King N. Doing Template Analysis: Evaluating an End of Life Care Service. Sage Research Methods Cases 2014:1-20 [FREE Full text] [doi: 10.4135/9781446273050135127555]
15. Brooks J, McCluskey S, Turley E, King N. The Utility of Template Analysis in Qualitative Psychology Research. Qual Res Psychol 2015 Apr 03;12(2):202-222 [FREE Full text] [doi: 10.1080/14780887.2014.955224] [Medline: 27499705]
16. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
17. 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]
18. Dossett LA, Kaji AH, Cochran A. SRQR and COREQ Reporting Guidelines for Qualitative Studies. JAMA Surg 2021 Sep 01;156(9):875-876. [doi: 10.1001/jamasurg.2021.0525] [Medline: 33825809]
19. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
20. Kostoff RN. Semiquantitative methods for research impact assessment. Technological Forecasting and Social Change 1993 Nov;44(3):231-244. [doi: 10.1016/0040-1625(93)90070-n]
21. Webster CS, McKillop A, Bennett W, Bagg W. A Qualitative and Semiquantitative Exploration of the Experience of a Rural and Regional Clinical Placement Programme. Med Sci Educ 2020 Jun;30(2):783-789 [FREE Full text] [doi: 10.1007/s40670-020-00949-6] [Medline: 34457735]
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
23. Zarbock A, Koyner JL, Hoste EAJ, Kellum JA. Update on Perioperative Acute Kidney Injury. Anesth Analg 2018 Nov;127(5):1236-1245. [doi: 10.1213/ANE.0000000000003741] [Medline: 30138176]
24. Schick AG, Gordon LA, Haka S. Information overload: A temporal approach. Accounting, Organizations and Society 1990 Jan;15(3):199-220. [doi: 10.1016/0361-3682(90)90005-f]
25. Cameron KA, Cohen ER, Hertz JR, Wayne DB, Mitra D, Barsuk JH. Barriers and Facilitators to Central Venous Catheter Insertion: A Qualitative Study. J Patient Saf 2021 Dec 01;17(8):e1296-e1306. [doi: 10.1097/PTS.0000000000000477] [Medline: 29543666]
26. Teichgräber UK, Benter T, Gebel M, Manns MP. A sonographically guided technique for central venous access. AJR Am J Roentgenol 1997 Sep;169(3):731-733. [doi: 10.2214/ajr.169.3.9275887] [Medline: 9275887]
27. Agarwal R, Karahanna E. Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. MIS Quarterly 2000 Dec;24(4):665. [doi: 10.2307/3250951]
28. Wachter SB, Agutter J, Syroid N, Drews F, Weinger MB, Westenskow D. The employment of an iterative design process to develop a pulmonary graphical display. J Am Med Inform Assoc 2003;10(4):363-372 [FREE Full text] [doi: 10.1197/jamia.M1207] [Medline: 12668693]
29. Ventilation Cockpit: A common user interface. Hamilton Medical. 2021. URL: https://www.hamilton-medical.com/ja/Solutions/Ventilation-Cockpit-user-interface.html [accessed 2021-11-30]
30. Neergaard MA, Olesen F, Andersen RS, Sondergaard J. Qualitative description - the poor cousin of health research? BMC Med Res Methodol 2009 Jul 16;9:52 [FREE Full text] [doi: 10.1186/1471-2288-9-52] [Medline: 19607668]
31. Carr LT. The strengths and weaknesses of quantitative and qualitative research: what method for nursing? J Adv Nurs 1994 Oct;20(4):716-721. [doi: 10.1046/j.1365-2648.1994.20040716.x] [Medline: 7822608]
32. Merry AF, Hannam JA, Webster CS, Edwards K, Torrie J, Frampton C, et al. Retesting the Hypothesis of a Clinical Randomized Controlled Trial in a Simulation Environment to Validate Anesthesia Simulation in Error Research (the VASER Study). Anesthesiology 2017 Mar;126(3):472-481 [FREE Full text] [doi: 10.1097/ALN.0000000000001514] [Medline: 28079566]
33. Tscholl DW, Rössler J, Said S, Kaserer A, Spahn DR, Nöthiger CB. Situation Awareness-Oriented Patient Monitoring with Visual Patient Technology: A Qualitative Review of the Primary Research. Sensors (Basel) 2020 Apr 09;20(7):2112 [FREE Full text] [doi: 10.3390/s20072112] [Medline: 32283625]

Edited by A Kushniruk; submitted 04.11.21; peer-reviewed by C Webster, M Hellaby; comments to author 25.11.21; revised version received 11.01.22; accepted 03.02.22; published 18.03.22.

Please cite as:
Akbas S, Said S, Roche TR, Nöthiger CB, Spahn DR, Tscholl DW, Bergauer L
User Perceptions of Different Vital Signs Monitor Modalities During High-Fidelity Simulation: Semiquantitative Analysis
JMIR Hum Factors 2022;9(1):e34677
URL: https://humanfactors.jmir.org/2022/1/e34677
doi:10.2196/34677
PMID:35119375

Original Paper

Clinician Perspectives on Unmet Needs for Mobile Technology Among Hospitalists: Workflow Analysis Based on Semistructured Interviews

April Savoy[1,2,3], PhD; Jason J Saleem[4], PhD; Barry C Barker[1], MA; Himalaya Patel[1], PhD; Areeba Kara[5,6], MS, MD

[1]Center for Health Information and Communication, Health Services Research and Development Service, Richard L Roudebush Veterans Affairs Medical Center, Indianapolis, IN, United States

[2]Purdue School of Engineering and Technology, Indiana University-Purdue University Indianapolis, Indianapolis, IN, United States

[3]Center for Health Services Research, Regenstrief Institute, Inc, Indianapolis, IN, United States

[4]Department of Industrial Engineering, James Breckenridge Speed School of Engineering, University of Louisville, Louisville, KY, United States

[5]Indiana University Health Physicians, Indianapolis, IN, United States

[6]Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
April Savoy, PhD
Center for Health Information and Communication
Health Services Research and Development Service
Richard L Roudebush Veterans Affairs Medical Center
1481 West 10th Street, Stop 11H
Indianapolis, IN, 46202
United States
Phone: 1 3172782194
Email: asavoy@iu.edu

Abstract

Background: The hospitalist workday is cognitively demanding and dominated by activities away from patients' bedsides. Although mobile technologies are offered as solutions, clinicians report lower expectations of mobile technology after actual use.

Objective: The purpose of this study is to better understand opportunities for integrating mobile technology and apps into hospitalists' workflows. We aim to identify difficult tasks and contextual factors that introduce inefficiencies and characterize hospitalists' perspectives on mobile technology and apps.

Methods: We conducted a workflow analysis based on semistructured interviews. At a Midwestern US medical center, we recruited physicians and nurse practitioners from hospitalist and inpatient teaching teams and internal medicine residents. Interviews focused on tasks perceived as frequent, redundant, and difficult. Additionally, participants were asked to describe opportunities for mobile technology interventions. We analyzed contributing factors, impacted workflows, and mobile app ideas.

Results: Over 3 months, we interviewed 12 hospitalists. Participants collectively identified chart reviews, orders, and documentation as the most frequent, redundant, and difficult tasks. Based on those tasks, the intake, discharge, and rounding workflows were characterized as difficult and inefficient. The difficulty was associated with a lack of access to electronic health records at the bedside. Contributing factors for inefficiencies were poor usability and inconsistent availability of health information technology combined with organizational policies. Participants thought mobile apps designed to improve team communications would be most beneficial. Based on our analysis, mobile apps focused on data entry and presentation supporting specific tasks should also be prioritized.

Conclusions: Based on our results, there are prioritized opportunities for mobile technology to decrease difficulty and increase the efficiency of hospitalists' workflows. Mobile technology and task-specific mobile apps with enhanced usability could decrease overreliance on hospitalists' memory and fragmentation of clinical tasks across locations. This study informs the design and implementation processes of future health information technologies to improve continuity in hospital-based medicine.

(JMIR Hum Factors 2022;9(1):e28783) doi:10.2196/28783

KEYWORDS

electronic health records; hospital medicine; user-computer interface; human-computer interaction; usability; mental workload; workflow analysis

Introduction

Electronic health record (EHR) systems aid documentation, information retrieval, and order creation. However, their lack of portability hampers effective support of communication between health care professionals and optimal access to patient information [1-3]. Such deficiencies contribute to task redundancies, constrain medical decisions at the point of care, and create inefficiencies that detract from valuable clinician-patient interactions [4-6]. These deficits are perhaps most impactful for hospitalists, a medical subspecialty focused on inpatient needs [7,8]. Multiple factors, including high patient acuity, ineffective health information technology (IT), hospital layouts, organizational policies, and interruptions, make hospitalists' workflow cognitively demanding and dominated by activities away from the patient's bedside (indirect care [9-15]).

As smartphones and tablets (mobile technology) became ubiquitous, they were proposed as one way to improve health IT. Physicians in emergency departments anticipated that these devices would improve workflow and physician-patient interactions [16], and in 2012, 87% of physicians were using smartphones and tablets in the workplace [17]. However, users in health care settings report lower expectations of mobile devices after actual use [17-19]. Most studies report experiences of physicians in training or those working in emergency departments. Less is known about the perceptions of hospitalists or their unique needs [20,21].

To improve care for Veteran patients, the US Department of Veterans Affairs (VA) Mobile Health Provider Program was launched in 2014. Through this program, over 12,000 iPads have been distributed at more than 60 VA sites. The program used a multiphase implementation strategy, focused on infrastructure updates, secure access to native mobile apps, and development of VA provider apps. However, adoption and use of the iPads and mobile apps among hospitalists has been low [22]. Our objective was to describe the needs and opportunities for mobile technology during the hospitalist workday. To characterize mobile technology that can synergistically support the workflow of hospitalists, we interviewed hospitalists to gain their perspectives on integrating mobile apps.

Methods

Overview

We conducted semistructured interviews guided by the Systems Engineering Initiative for Patient Safety (SEIPS) framework [5,23]. This framework, consisting of five factors (people, environment, tasks, tools, and organization) and their interactions, can be used to describe how health care providers' work systems impact workflows and outcomes [5]. Our interviews focused on tasks from multiple workflows to obtain in-depth information about related frequencies, redundancies, difficulties, and mobile apps. Our analysis aimed to characterize contributing work system factors, multiple impacted workflows, and participants' ideas for mobile app interventions.

Participants and Setting

The study was conducted at a 200-bed urban teaching hospital operated by VA in Indiana. This hospital offered iPad tablets and introductory training to its health care providers. We sought approximately 12 participants to increase the likelihood of thematic saturation [24,25]. We sought participants who practiced according to the hospitalist model of care because they may face overlapping workflow challenges. Physicians and nurse practitioners from hospitalist and inpatient teaching teams and second- or third-year internal medicine residents were eligible to participate [26,27]. Eligible participants were contacted via email, and a nonfiction book was offered for participation. This study was approved and overseen by the Institutional Review Board at Indiana University (#1608865326) and the Research and Development committee at Richard L. Roudebush VA Medical Center.

Semistructured Interviews

Data Collection

Semistructured interviews were designed as 45-minute sessions (Multimedia Appendix 1). Participants were asked to describe their primary roles and information-intensive tasks. They were then asked to identify tasks that were frequent, redundant, and difficult, and to explain their choices [28,29]. Definitions were reviewed with participants as follows:

- Information-intensive tasks: require reading, writing, or sharing information (eg, chart review)
- Frequent: performed often or for each patient (eg, looking up patients' contact information or reviewing discharge summaries)
- Redundant: done repeatedly that should only be done once or not at all (eg, repetitious log-ins or clicks to access required information)
- Difficult: require uninterrupted time and attention (eg, reviewing labs or determining trends in vitals)

Participants were asked to describe the context of each task with a focus on work system factors [5]. Interviewers diagrammed discussions as participants spoke. Participants completed a demographic survey, including reporting use of self- and work-furnished mobile technology. Demographics and field notes were collected on paper and scanned. Interviews were audio recorded and transcribed.

Workflow Analysis

For each task identified as frequent, redundant, or difficult, we open-coded participants' responses and organized these codes within the five work system factors. Next, we analyzed the impact those tasks had on workflows [30,31]. Lastly, we analyzed participants' responses to mobile technology to identify and describe types of potential mobile app interventions.

Using a hybrid deductive-inductive approach, we iteratively developed a code book with sections and codes to aid each type of analysis [32]. We used a deductive approach to identify relevant work system factors and an inductive approach to describe workflow effects and potential mobile app interventions. One analyst created the preliminary code book based on the SEIPS work system factors and open coding of three transcripts. Using this preliminary code book, four additional analysts reviewed another set of three transcripts. The team discussed and refined codes. With the revised code book, four analysts worked in pairs to code the remaining transcripts, which were randomly assigned. After coding each transcript independently, coding partners reviewed transcripts line by line, resolving discrepancies through consensus meetings. If new codes emerged during coding, they were retroactively applied to previously coded transcripts. Each analyst wrote memos for prominent codes; then, analyst pairs conducted consensus meetings. After these meetings, an analyst selected the most frequent open codes, linked the most frequently co-occurring codes for each, and prepared a narrative summary with supporting quotations. Coding, memo writing, and content analysis were performed using Excel (Version 2016, Microsoft Corporation).

Results

Participants

Over 3 months, we interviewed 12 participants: 9 staff physicians, 1 resident physician, and 2 nurse practitioners. Including residency, experience ranged from 0.9 to 37 years (mean 11.7, median 8.5); experience in the present organization was similar (0.5-37 years; mean 10.8, median 8.5). A total of 11 (92%) reported using mobile technology at work, including both personally owned and work-furnished devices. Only 3 (25%) reported using iPad tablets at work. Nurse practitioners worked on the hospitalist team, while staff physicians rotated between the nonteaching hospitalist team and inpatient teaching teams. Patient load was estimated as ranging from an average of 10 to 15 patients per day.

In the following section, we first present participants' perceptions of specific tasks that were perceived as frequent, redundant, and difficult. Next, we describe the workflows that were perceived to be most impacted by these tasks. Finally, we present participants' perceptions of mobile technology and potential mobile app interventions.

Frequent, Redundant, and Difficult Tasks

Chart reviews, orders, and documentation were identified as the most frequent, redundant, and difficult tasks.

Chart Reviews

Participants described that chart reviews (re)established the patient's trajectory, which included viewing patient history, recent notes, laboratory results, and vitals. Participants reported conducting a summary review for every patient throughout the day to monitor progress, orders, procedures, and test results, and estimated spending 30 minutes to review the chart of a new patient. Some participants noted that their initial reviews were completed at the beginning of the day or before their shift. This need for on-demand continuity was contrasted with fragmentation of records in EHRs and the multistep methods for accessing them. Participants characterized chart review as redundant because of intermittent updates without notification, resulting in checking either too little or too often. Information copied in workrooms and carried to patients could be outdated upon reaching patients, or effort could be wasted looking for information that had not yet arrived.

Orders

Participants described writing orders multiple times a day using computers. Orders included lab tests, consultations, and prescriptions. The institution currently requires electronic entry of all orders. Participants described the lack of (bedside) computers, not necessarily the need for complex thought, was what made ordering difficult and inefficient. Perception of ordering was also negatively affected by poor EHR usability. The organization of orders in the EHR was thought to be unclear, decreasing the discoverability of specific orders. Participants gave examples of order forms for similar procedures that were found in different branches of the menu. This poor organization of order forms was described as increasing difficulty by limiting application of knowledge between orders—finding and writing one type of order did not necessarily make it easier to find or write other types.

Documentation

Documentation was reported as one of the most labor- and time-intensive tasks. It included documenting a variety of information, including histories and physicals, visit notes, daily note, and discharge summaries. As with chart review, fragmentation of information in EHRs meant writing notes frequently, even when new notes were similar to previous notes. Participants described EHR documentation as a constant process consuming a considerable portion of the day. Patient load was estimated as ranging from an average of 10 to 15 patients per day. With that, participants estimated that documentation time averaged 30 to 45 minutes per patient. Participants described documentation as redundant, as they and their trainees were required to write notes for the same patients.

Table 1 summarizes the contributing work system factors for each task perceived as frequent, redundant, and difficult with some illustrative quotes from participants.

Table 1. Frequent, redundant, and difficult tasks identified by participants and derived contributing work system factors.

Task	Contributing work system factor(s)	Representative quotation
Chart review[a]	• People: Extent of reliance on electronic records varied between participants • Environment: Electronic chart was not accessible at bedsides • Tools/technology: EHR[b] did not push notifications of important changes • Tasks: Patients with more status changes needed more frequent review • Organization: Multifactor authentication was required before every EHR session	"Ideally, you would like to be able to harvest that information in the room with the patient by handheld device so that if memory fails and patients have questions, you can use that to help answer their questions. Mostly, I do that from memory now."
Orders[c]	• People: Preferences varied in when to start and when to submit orders • Environment: Electronic ordering was not accessible at bedsides • Tools/technology: Finding the right order form in the EHR was difficult • Tasks: Orders depended on having the most up-to-date patient information • Organization: All orders had to be made through the EHR	"There are multiple clicks to get to different boxes, lots of pop-ups that you have to go through...the computer system itself adds considerably to the amount of time that we take and takes away from our patient care"
Documentation[d]	• People: Content of attendings' notes depended on the content of their residents' notes • Environment: EHR was not accessible at bedsides • Tools/technology: Authoring notes in the EHR sometimes involved copying forward text from older notes • Tasks: These were sometimes based on a single encounter, and other times more longitudinal (eg, discharge summaries) • Organization: Facility required a series of documentation and ordering steps before discharge	"I think documentation is by far the thing that takes us the longest— documentation for sure."

[a]Chart review: going through patient information and history.

[b]EHR: electronic health record.

[c]Orders: services like lab tests and referral.

[d]Documentation: summarizing encounters, making or changing care plans, and adding to patient information.

Impact of Tasks: Inefficient and Difficult Hospitalists' Workflows

Participants characterized admit, discharge, and rounding workflows as difficult or inefficient.

Intake

The admit workflow was reported to be time-consuming:

> *It takes 1.5-2 hours to do an admission from start to finish,...entails chart review, seeing the patient, putting in orders, reviewing things, and doing the history and physical.*

Difficulty of completing tasks seemingly increased as the workday progressed. Often, patients' care was distributed across multiple health care systems. In those cases, admitting was described as involving retrieving both internal and external records. At best, external records were retrieved electronically (eg, from a health information exchange). Otherwise, retrieving outside records involved making telephone calls and reviewing scanned records. Some participants relied on residents:

> *I usually have learners helping take care of some other tasks but without learners sometimes it [compiling patients' histories] just doesn't happen.*

Discharge

Most participants noted efforts to complete discharges by lunchtime:

> *...it's usually like a 4-5 hour process. It's challenging to discharge patients in the afternoon, because there's just too much to get done. It's cumbersome...*

Due to documentation demands, participants described these workflows as redundant and time-consuming. Discharges involved data retrieval that depended on the length of the stay and much documentation. Several notes needed to be written, and among those notes, a large amount of information was duplicated:

> *...So discharge note, anticipated note, discharge instruction, discharge summary, medical reconciliation, pharmacy output...we can clump together to save time...*

These characteristics related to admit and discharge workflows increased participants' time in workrooms because access to their desktops were required to complete notes.

Rounds

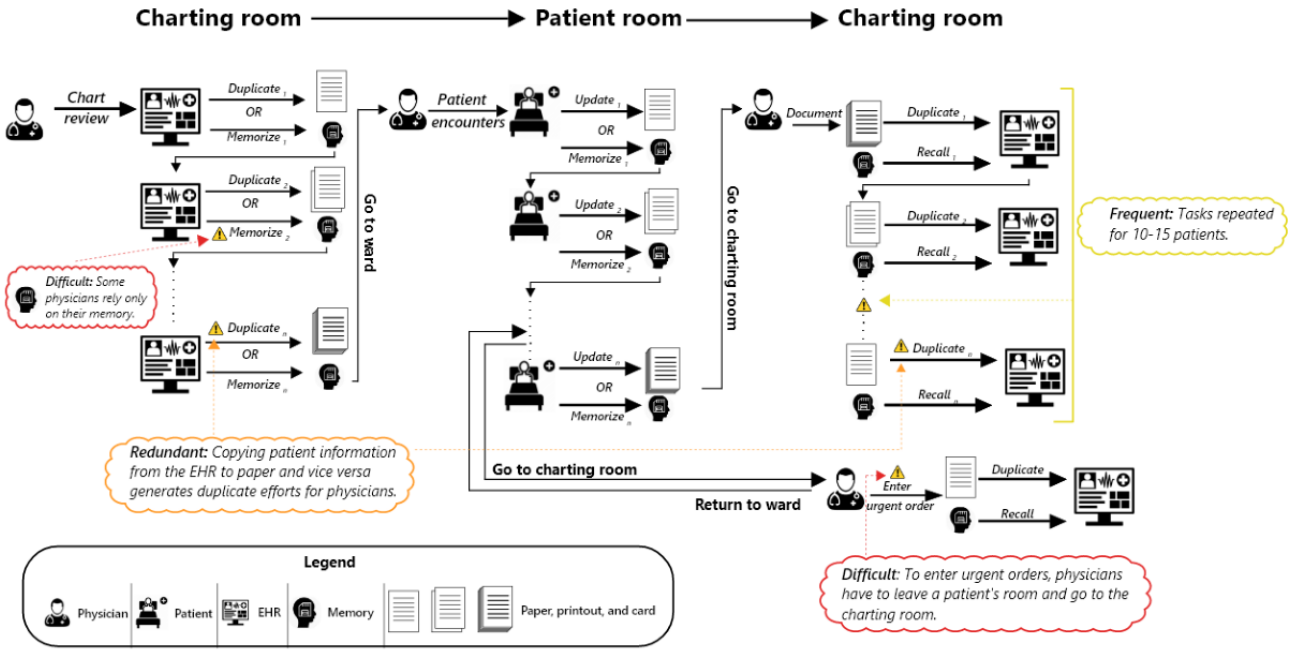
Rounding was identified as an inefficient workflow. Participants reported seeing 10 to 18 patients during rounding. For each patient, participants documented history and physical notes in

the EHR. Afterward, they duplicated the text on paper cards to support rounding. Otherwise, the information was not readily available. Printouts were wasteful because page counts averaged 6 pages per patient. Unlike computers and the EHR, paper notes fit in their pockets and were easily accessible for review and modification:

> *Bedside computers are not at the bedside. We don't really have access to computers that work very well other than those in our team rooms.*

Based on participant interviews, Figure 1 illustrates a snapshot of the participants' description of difficult and inefficient workflows that stem from frequent, redundant, and difficult tasks.

Figure 1. Snapshot of difficult and inefficient elements in hospitalists' workflows. Hospitalists start in the charting room and conduct chart reviews for all patients they will visit. For each patient, hospitalists must duplicate information from the EHR on index cards or printouts to support review at bedside. After completing cards, hospitalists take all the cards to the ward where patients are located. Hospitalists find the appropriate card for each patient encounter and update the card with new patient information related to status, orders, and plans. Hospitalists move from one patient to the next, repeating those steps. After the last patient encounter, hospitalists go back to the charting room to enter the information from the cards into the EHR. This entire workflow is done multiple times a day. EHR: electronic health record.



Participants' Perspectives on Opportunities for Mobile App Interventions

Participants expected mobile technology to decrease task completion time; however, they noted that neither rapid access nor documentation of information was supported by current mobile apps. Usability issues were also noted, highlighting the misalignment between expected and actual functionality. One participant said:

> *I had an iPad for a while here when I was in the pain clinic, but I didn't use it. I couldn't do controlled substances refills on it, and that was all that I ordered in the pain clinic. So I turned it back in...*

As a result, participants expressed that they viewed paper and their brains as "surrogate mobile devices." Paper to-do lists were described as repositories of patient information and task trackers. Paper was perceived as more reliable than mobile technology. Alternatively, some participants described heavy reliance on their memory. As one participant noted, pointing to his head, "my technology is up here."

When discussing ideas for mobile technology, participants prioritized portability, reduction in task time completion and task completion at bedside (Textbox 1). Three representative examples of useful mobile apps emerged. First, participants said a mobile device like an iPad would help patient-provider communication and entering orders at bedside. Second, participants described a note-taking app that had sharing features and stored nurses contact information. Lastly, participants proposed an app for electronic consents.

Textbox 1. Participant quotes describing potential mobile app solutions.

Patient-provider communication

"...instead of telling patients, actually giving them a visual as you are rounding will make them help feel more involved in their own care."

Team communication

"...[communication] breakdown occurs when we're calling nurses...if we just had the correct number in the first place, we wouldn't have to go through talking to multiple people."

eConsent

"...when you go buy a coffee and doughnuts, you know how you can just sign on the iPad; having that same setup for consent may work well."

Discussion

Summary

Provider-EHR interactions in inpatient care have contributed to increased workflow inefficiencies, reduced time for provider-patient interactions, and increased cognitive burden among hospitalists. Our findings provide a better understanding of the misalignment between hospitalists' needs and expectations of available mobile technology and evidence of hospitalists' cognitively intense workflows. In this section, we discuss our findings and implications for implementation of future mobile technology interventions for hospitalists.

Influential Contextual Factors

Despite the need for mobile access to patient information, mobile technology was not widely adopted. Difficult and inefficient tasks were predominately related to provider-EHR interactions because access to EHRs was not consistent at bedside [33]. We associated hospitalists' unmet needs with one or more of the following SEIPS factors: tasks, tools and technology, and environment (location). For example, participants had to travel across three floors to complete workflows and clinical tasks that required chart review, patient encounters, and documentation. These dynamics influenced the perceived difficulty, frequency, and redundancy among workflows and clinical tasks. Although mobile technology was available, usability issues related to existing mobile apps prevented their use, increasing participants' reliance on index cards and printouts. Thus, contextual factors influenced the need for mobile technology, but misalignments of hospitalists' expectations and mobile device functionality limited the adoption and use of existing mobile apps. This finding demonstrates the critical importance of integrating workflow analysis into the design process of mobile technology interventions; the result of this analysis identifies unmet needs and unintended consequences.

Implications of Cognitive Workload and Burden

The three information-intensive tasks (chart review, ordering, and documentation) identified as frequent, redundant, and difficult were prone to an overreliance on hospitalists' memory, including working memory [34]. A major contributor to this overreliance was the lack of mobile technology that supported chart review or order entry needed at bedside. Classifying tasks as frequent and redundant were easy for participants. These tasks were often described as inefficient and sources of hospitalists' frustration. Identifying tasks as difficult caused participants to think of their tasks in a new way. Echoed throughout our data collection, tasks were not difficult due to hospitalists' lack of knowledge or training to identify treatment plans, make clinical decisions, or perform clinical procedures. Rather, difficulty was defined and associated with the lack of support and access to usable technology required to review and enter information at bedside. According to existing cognition literature [35-37], these workflow aspects required participants to change location frequently, which increased the likelihood of interruptions and limited information recall (ie, cognitive slips and mistakes). This can be linked to incomplete documentation, communication breakdowns, and delays in care alluded to in participant interviews. Paper-based work-arounds were associated with processing orders and notes together in one sitting (ie, batch processing), not individually at the time of each decision. Batch processing has been associated with delayed team communication, delayed discharges, and shift limit violations [38]. Although the terms *cognitive burden* and *mental overload* were not specifically mentioned in interviews, these were clear outcomes for hospitalists based on our analysis. Cognitive burden can decrease resilience, situation awareness, and subsequently patient safety [36,39-41].

Potential for Task-Specific Mobile Apps

Hospitalists thought task-specific apps would be most helpful. Their primary goals were to reduce inefficiencies or difficulties with orders, discharge, consent, and team communication. Hospitalists' focus on individual tasks indicates a need to shift design goals of mobile apps that focused on granting access to the entire EHR via consistent user interfaces (eg, mobile version of EHR desktop interfaces). Participants stressed the need for task-specific apps that highlighted fast, focused technology interactions when away from the charting room. The design of mobile apps should be based on the objective and use of the paper or cards currently used for hospitalists' mobile workflows, including quick review and documentation of prioritized patient information. For example, apps should present customized views of patient health status or trends. In addition, apps can support bedside order entry with *smart* templates that use automation or dictation to optimize data entry without keyboards. Based on our findings, mobile apps designed to support the iterative nature of hospitalists' workflows or rounds by providing a means to review charts and document at bedside may reduce the need for batch processing before and after patient encounters. Thus, current workflows would be streamlined, decreasing the redundancy and difficulty illustrated in workflows characterized in our study.

There are several task-specific apps that are being trialed and should be monitored for success. Since the completion of our

study, the VA's Office of Connected Care is working to achieve greater understanding of provider preferences for mobile technology and task-specific apps. Providers currently have access to a variety of task-specific apps for mobile computing through the VA App Store. For example, the Image Viewing Solution is an app to access diagnostic-grade images. Annie App for Clinicians allows providers to assign disease-specific protocols to their patients. Several other task-specific apps are in development to meet VA providers' needs.

Limitations

This workflow analysis was limited by a relatively small sample in one health care facility. VA is the nation's largest integrated health care system. Therefore, participant perspectives of hospitalists' workflows and mobile technology may be broadly relevant to other health care systems. For example, initial deployment of mobile technology, without a variety of task-specific clinical apps readily available contributed to the low adoption of mobile technology [18,19]. By using informal definitions of frequent, redundant, and difficult, these concepts may have overlapped to some degree. We did not explore differences associated with career stage (eg, early, middle, and late). Our findings demonstrated the influence of contextual factors; future studies should further explore interactions between technology use, interruptions, and geographic cohorting across multiple facilities [35,42-44].

Conclusion

Based on our results, there are opportunities for mobile technology to decrease the difficulty and increase the efficiency of hospitalists' workflows. Mobile technology and task-specific mobile apps with enhanced usability have the potential to decrease overreliance on hospitalists' memory and fragmentation of clinical tasks across locations that exist with current health IT and hospital environments. Task-specific apps that aim to reduce redundancies or excessive administrative work related to admissions, orders, and discharges were prioritized by hospitalists. Human factors engineering approaches are needed to identify hospitalists' requirements for mobile technology to address issues with information management and recall during rounds. Extending beyond hardware features, a better understanding of direct and contextual factors of mobile information needs is required to develop mobile apps that can support hospitalists' workflows. This will be influential in initial and sustained adoption of future mobile technology and apps.

Acknowledgments

This work was supported by a pilot grant (PPO 15-401; AS) and a Center of Innovation grant (CIN 13-416, M Weiner), both from the United States Department of Veterans Affairs Health Services Research and Development. AS is supported in part by the following grants: KL2TR002530 (A Carroll, PI), and UL1TR002529 (A Shekhar, PI) from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award.

Authors' Contributions

AS proposed the study and secured funding. BCB wrote the interview guide with input from AS, HP, and research assistants (Diana Natividad and Rachel Dismore). BCB led interviews and analysis. AS drafted the manuscript. Anna Mathew created the workflow illustration. All authors interpreted the findings, made critical revisions, and approved the published manuscript; all authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 16 KB - humanfactors_v9i1e28783_app1.docx]

References

1. Saleem JJ, Russ AL, Neddo A, Blades PT, Doebbeling BN, Foresman BH. Paper persistence, workarounds, and communication breakdowns in computerized consultation management. Int J Med Inform 2011 Jul;80(7):466-479. [doi: 10.1016/j.ijmedinf.2011.03.016] [Medline: 21530383]
2. Lowry SZ, Ramaiah M, Patterson ES, Brick D, Gurses AP, Ozok A, et al. Integrating Electronic Health Records Into Clinical Workflow: An Application of Human Factors Modeling Methods to Ambulatory Care. Washington, DC: US Department of Commerce; 2014.
3. Russ AL, Saleem JJ, Justice CF, Woodward-Hagg H, Woodbridge PA, Doebbeling BN. Electronic health information in use: characteristics that support employee workflow and patient care. Health Informatics J 2010 Dec;16(4):287-305 [FREE Full text] [doi: 10.1177/1460458210365981] [Medline: 21216808]

4. Kim MO, Coiera E, Magrabi F. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. J Am Med Inform Assoc 2017 Mar 01;24(2):246-250 [FREE Full text] [doi: 10.1093/jamia/ocw154] [Medline: 28011595]
5. Carayon P, Wetterneck TB, Rivera-Rodriguez AJ, Hundt AS, Hoonakker P, Holden R, et al. Human factors systems approach to healthcare quality and patient safety. Appl Ergon 2014 Jan;45(1):14-25 [FREE Full text] [doi: 10.1016/j.apergo.2013.04.023] [Medline: 23845724]
6. Erickson SM, Rockwern B, Koltov M, McLean RM, Medical Practice and Quality Committee of the American College of Physicians. Putting patients first by reducing administrative tasks in health care: a position paper of the American College of Physicians. Ann Intern Med 2017 May 02;166(9):659-661 [FREE Full text] [doi: 10.7326/M16-2697] [Medline: 28346948]
7. Wachter RM, Goldman L. Zero to 50,000 - The 20th Anniversary of the Hospitalist. N Engl J Med 2016 Sep 15;375(11):1009-1011. [doi: 10.1056/NEJMp1607958] [Medline: 27508924]
8. 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]
9. Salahuddin L, Ismail Z. Classification of antecedents towards safety use of health information technology: a systematic review. Int J Med Inform 2015 Nov;84(11):877-891. [doi: 10.1016/j.ijmedinf.2015.07.004] [Medline: 26238706]
10. Weigl M, Müller A, Vincent C, Angerer P, Sevdalis N. The association of workflow interruptions and hospital doctors' workload: a prospective observational study. BMJ Qual Saf 2012 May;21(5):399-407. [doi: 10.1136/bmjqs-2011-000188] [Medline: 22190539]
11. 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]
12. Gurvich I, O'Leary KJ, Wang L, Van Mieghem JA. Collaboration, interruptions, and changeover times: workflow model and empirical study of hospitalist charting. Manufacturing Serv Operations Manage 2020 Jul;22(4):754-774. [doi: 10.1287/msom.2019.0771]
13. Baumann LA, Baker J, Elshaug AG. The impact of electronic health record systems on clinical documentation times: a systematic review. Health Policy 2018 Aug;122(8):827-836. [doi: 10.1016/j.healthpol.2018.05.014] [Medline: 29895467]
14. Tsai CY, Pancoast P, Duguid M, Tsai C. Hospital rounding--EHR's impact. Int J Health Care Qual Assur 2014;27(7):605-615. [doi: 10.1108/ijhcqa-07-2013-0090] [Medline: 25252566]
15. Assis-Hassid S, Grosz BJ, Zimlichman E, Rozenblum R, Bates DW. Assessing EHR use during hospital morning rounds: a multi-faceted study. PLoS One 2019;14(2):e0212816 [FREE Full text] [doi: 10.1371/journal.pone.0212816] [Medline: 30802267]
16. Rao AS, Adam TJ, Gensinger R, Westra BL. Study of the factors that promoted the implementation of electronic medical record on iPads at two emergency departments. AMIA Annu Symp Proc 2012;2012:744-752 [FREE Full text] [Medline: 23304348]
17. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
18. Walsh C, Stetson P. EHR on the move: resident physician perceptions of iPads and the clinical workflow. AMIA Annu Symp Proc 2012;2012:1422-1430 [FREE Full text] [Medline: 23304422]
19. Horng S, Goss FR, Chen RS, Nathanson LA. Prospective pilot study of a tablet computer in an Emergency Department. Int J Med Inform 2012 May;81(5):314-319 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.12.007] [Medline: 22226927]
20. O'Leary KJ, Liebovitz DM, Baker DW. How hospitalists spend their time: insights on efficiency and safety. J Hosp Med 2006 Mar;1(2):88-93. [doi: 10.1002/jhm.88] [Medline: 17219478]
21. Tipping MD, Forth VE, O'Leary KJ, Malkenson DM, Magill DB, Englert K, et al. Where did the day go?--a time-motion study of hospitalists. J Hosp Med 2010;5(6):323-328. [doi: 10.1002/jhm.790] [Medline: 20803669]
22. Saleem JJ, Savoy A, Etherton G, Herout J. Investigating the need for clinicians to use tablet computers with a newly envisioned electronic health record. Int J Med Inform 2018 Feb;110:25-30. [doi: 10.1016/j.ijmedinf.2017.11.013] [Medline: 29331252]
23. Holden RJ, Carayon P, Gurses AP, Hoonakker P, Hundt AS, Ozok AA, et al. SEIPS 2.0: a human factors framework for studying and improving the work of healthcare professionals and patients. Ergonomics 2013;56(11):1669-1686 [FREE Full text] [doi: 10.1080/00140139.2013.838643] [Medline: 24088063]
24. Crandall B, Klein GA, Hoffman RR. Working Minds: A Practitioner's Guide to Cognitive Task Analysis. Cambridge, MA: The MIT Press; 2006.
25. Guest G, Bunce A, Johnson L. How many interviews are enough? Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]
26. Pantilat S. What is a hospitalist? The Hospitalist. 2006 Feb. URL: https://www.the-hospitalist.org/hospitalist/article/123072/what-hospitalist [accessed 2021-12-01]
27. Cowan MJ, Shapiro M, Hays RD, Afifi A, Vazirani S, Ward CR, et al. The effect of a multidisciplinary hospitalist/physician and advanced practice nurse collaboration on hospital costs. J Nurs Adm 2006 Feb;36(2):79-85. [doi: 10.1097/00005110-200602000-00006] [Medline: 16528149]

28. Morgeson FP. Job analysis methods. In: Rogelberg SG, editor. The SAGE encyclopedia of industrial and organizational psychology. Thousand Oaks, CA: SAGE Publications; 2017:769-771.
29. Morgeson FP, Brannick MT, Levine EL. Job and Work Analysis: Methods, Research, and Applications for Human Resource Management. Los Angeles, CA: SAGE Publications; 2019.
30. Unertl KM, Weinger MB, Johnson KB, Lorenzi NM. Describing and modeling workflow and information flow in chronic disease care. J Am Med Inform Assoc 2009;16(6):826-836 [FREE Full text] [doi: 10.1197/jamia.M3000] [Medline: 19717802]
31. Unertl KM, Novak LL, Johnson KB, Lorenzi NM. Traversing the many paths of workflow research: developing a conceptual framework of workflow terminology through a systematic literature review. J Am Med Inform Assoc 2010;17(3):265-273 [FREE Full text] [doi: 10.1136/jamia.2010.004333] [Medline: 20442143]
32. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. Psychiatry Res 2019 Oct;280:112516 [FREE Full text] [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
33. Clynch N, Kellett J. Medical documentation: part of the solution, or part of the problem? A narrative review of the literature on the time spent on and value of medical documentation. Int J Med Inform 2015 Apr;84(4):221-228. [doi: 10.1016/j.ijmedinf.2014.12.001] [Medline: 25547194]
34. Cowan N. The magical mystery four: how is working memory capacity limited, and why? Curr Dir Psychol Sci 2010 Feb 01;19(1):51-57 [FREE Full text] [doi: 10.1177/0963721409359277] [Medline: 20445769]
35. Zheng K, Haftel HM, Hirschl RB, O'Reilly M, Hanauer DA. Quantifying the impact of health IT implementations on clinical workflow: a new methodological perspective. J Am Med Inform Assoc 2010;17(4):454-461 [FREE Full text] [doi: 10.1136/jamia.2010.004440] [Medline: 20595314]
36. Radvansky GA, Krawietz SA, Tamplin AK. Walking through doorways causes forgetting: further explorations. Q J Exp Psychol (Hove) 2011 Aug;64(8):1632-1645. [doi: 10.1080/17470218.2011.571267] [Medline: 21563019]
37. Murre JMJ, Dros J. Replication and analysis of Ebbinghaus' forgetting curve. PLoS One 2015;10(7):e0120644 [FREE Full text] [doi: 10.1371/journal.pone.0120644] [Medline: 26148023]
38. Calderon AS, Blackmore CC, Williams BL, Chawla KP, Nelson-Peterson DL, Ingraham MD, et al. Transforming ward rounds through rounding-in-flow. J Grad Med Educ 2014 Dec;6(4):750-755 [FREE Full text] [doi: 10.4300/JGME-D-13-00324.1] [Medline: 26140131]
39. Farrell LJ, Du S, Steege LM, Cartmill RS, Wiegmann DA, Wetterneck TB, et al. Understanding cognitive requirements for EHR design for primary care teams. Proc Int Symp Hum Factors Ergonomics Health Care 2017 May 15;6(1):15-16. [doi: 10.1177/2327857917061005]
40. Belden JL, Patel J, Lowrance N, Plaisant C, Koopman R, Moore J, et al. Inspired EHRs: Designing for Clinicians. Columbia, MO: University of Missouri; 2014. URL: http://inspiredehrs.org/ [accessed 2021-12-01]
41. Singh H, Giardina TD, Petersen LA, Smith MW, Paul LW, Dismukes K, et al. Exploring situational awareness in diagnostic errors in primary care. BMJ Qual Saf 2012 Jan;21(1):30-38 [FREE Full text] [doi: 10.1136/bmjqs-2011-000310] [Medline: 21890757]
42. Kara A, Flanagan ME, Gruber R, Lane KA, Bo N, Kroenke K, et al. A time motion study evaluating the impact of geographic cohorting of hospitalists. J Hosp Med 2020 Jun;15(6):338-344. [doi: 10.12788/jhm.3339] [Medline: 31891555]
43. Lopetegui M, Yen PY, Lai AM, Embi PJ, Payne PRO. Time Capture Tool (TimeCaT): development of a comprehensive application to support data capture for Time Motion Studies. AMIA Annu Symp Proc 2012;2012:596-605 [FREE Full text] [Medline: 23304332]
44. Ballermann MA, Shaw NT, Mayes DC, Gibney RTN, Westbrook JI. Validation of the Work Observation Method By Activity Timing (WOMBAT) method of conducting time-motion observations in critical care settings: an observational study. BMC Med Inform Decis Mak 2011 May 17;11:32 [FREE Full text] [doi: 10.1186/1472-6947-11-32] [Medline: 21586166]

Abbreviations

EHR: electronic health record
IT: information technology
SEIPS: Systems Engineering Initiative for Patient Safety
VA: Veterans Affairs

Edited by A Kushniruk; submitted 16.03.21; peer-reviewed by D Scherer, J Kellett, T Koritala; comments to author 10.05.21; revised version received 01.06.21; accepted 10.10.21; published 04.01.22.

Please cite as:
Savoy A, Saleem JJ, Barker BC, Patel H, Kara A
Clinician Perspectives on Unmet Needs for Mobile Technology Among Hospitalists: Workflow Analysis Based on Semistructured Interviews
JMIR Hum Factors 2022;9(1):e28783
URL: https://humanfactors.jmir.org/2022/1/e28783
doi:10.2196/28783
PMID:34643530

Original Paper

Electronic Records With Tablets at the Point of Care in an Internal Medicine Unit: Before-After Time Motion Study

Montserrat Pérez-Martí[1,2], RN, MSc; Lina Casadó-Marín[3], RN, BS, MSc, PhD; Abraham Guillén-Villar[4], BS, BSocSc, MSc, MEd

[1]Department of Information Technology, Consorci Sanitari Alt Penedès-Garraf, Vilafranca del Penedès, Spain

[2]Department of Anthropology and Communication, Universitat Rovira i Virgili, Tarragona, Spain

[3]Department of Nursing, Universitat Rovira i Virgili, Tarragona, Spain

[4]Tarragona, Spain

Corresponding Author:
Montserrat Pérez-Martí, RN, MSc
Department of Information Technology
Consorci Sanitari Alt Penedès-Garraf
c/Espirall 61
Vilafranca del Penedès, 08720
Spain
Phone: 34 938194000 ext 7445
Email: mperezm@csapg.cat

Abstract

Background: There are many benefits of nursing professionals being able to consult electronic health records (EHRs) at the point of care. It promotes quality and patient security, communication, continuity of care, and time dedicated to records.

Objective: The aim of this study was to evaluate whether making EHRs available at the point of care with tablets reduces nurses' time spent on records compared with the current system. The analysis included sociodemographic and qualitative variables, time spent per patient, and work shift. This time difference can be used for direct patient care.

Methods: A before-after time motion study was carried out in the internal medicine unit. There was a total of 130 observations of 2 hours to 3 hours in duration of complete patient records that were carried out at the beginning of the nurses' work shifts. We calculated the time dedicated to measuring vital signs, patient evaluation, and EHR recording. The main variable was time spent per patient.

Results: The average time spent per patient (total time/patients admitted) was lower with the tablet group (mean 4.22, SD 0.14 minutes) than with the control group (mean 4.66, SD 0.12 minutes); there were statistically significant differences (W=3.20, *P*=.001) and a low effect (d=.44) between groups. The tablet group saved an average of 0.44 (SD 0.13) minutes per patient. Similar results were obtained for the afternoon shift, which saved an average of 0.60 (SD 0.15) minutes per patient (t_{34}=3.82, *P*=.01) and high effect (d=.77). However, although there was a mean difference of 0.26 (SD 0.22) minutes per patient for the night shift, this was not statistically significant (t_{29}=1.16, *P*=.25). The "nonparticipating" average age was higher (49.57, SD 2.92 years) compared with the "afternoon shift participants" and "night shift participants" (*P*=.007). "Nonparticipants" of the night shift had a worse perception of the project.

Conclusions: This investigation determined that, with EHRs at the point of care, the time spent for registration by the nursing staff decreases, because of reduced movements and avoiding data transcription. It eliminates unnecessary work that does not add value, and therefore, care is improved. So, we think EHRs at the point of care should be the future or natural method for nursing to undertake. However, variables that could have a negative effect include age, night shift, and nurses' perceptions. Therefore, it is proposed that training in the different work platforms and the participation of nurses are fundamental axes that any institution should consider before their implementation.

(JMIR Hum Factors 2022;9(1):e30512) doi:10.2196/30512

KEYWORDS

electronic health records; nursing; computer handheld; equipment and supplies (devices tablets mobile phones, devices and technologies); workflow

Introduction

Background

Nurses represent the category with the highest number of contracted professionals in the health care workforce and therefore the largest group of users of electronic health records (EHRs; in the countries where they are implemented) [1-3]. They use them as a primary tool for documenting, synthesizing, and communicating patient data. Therefore, introducing electronic devices into different work areas has a great impact on this group [2].

It is essential that the nurses are involved and committed to completing EHRs. Their use should be guided by nurses, as it is common for nursing professionals not to participate in their development [2,4-7]. EHRs should be usable for nurses and relevant for their practice [2,8].

Concepts such as usability, utility, efficiency in the context of the users' use, and trust in technology are key elements for nurses to accept this innovation [8]. Therefore, we need to know the contributions or opportunities that they offer us, as well as their limitations.

Bibliography authors frequently highlight the opportunity offered by EHRs for developing new instruments that improve the quality and efficiency of care, such as standardized care plans, checklists, and decision support systems. Standardized care plans also prioritize the need to make the profession visible by offering results that demonstrate the effectiveness of care [2,5-7,9-13].

However, some authors state that they can act as a limiting agent for care, reducing critical thinking, clinical judgment, and basic nursing skills [5,10,14,15].

Another category that is important for the profession is saving time and communication. Some authors argue that EHRs decrease the time spent recording, because they have facilitating instruments such as copy-and-paste or drop-down menus with content standardization. They also promote access to information, thus improving inter- and multidisciplinary communication and consequently improving continuity of care [8-10,12,15].

Nevertheless, limiting agents in this category are also of equal importance. There is a greater volume of data available from any device, and finding information relevant to health care practice is not always easy. The user is forced to navigate the entire system, opening multiple screens; or using different software, with duplicated information, to obtain an overview of the patient's condition. The user is also required to enter a password repeatedly for the different programs, or the software has poor performance; there is an excess of mandatory information that needs to be entered; there are frequent interruptions; computers at the workstations are constantly busy, unavailable, or shared; or the setup has unfavorable ergonomics that deter nurses from trying to obtain information during the person's point of care. All these limiting agents describe a poor or poorly developed system that can cause interruptions in workflows in which the nurse has to perform more steps to carry out a nursing activity than would be necessary. More time spent on EHRs, time that adds no value, can lead to unnecessary delays in patient care. Another time-consuming cause that authors have found is a high percentage of transcription on paper of patient data, resulting in delays in patient information arriving to other professionals and increasing the possibility of error. This is experienced by nurses as work overload, a time limiter, and a barrier to communication. As a result, disruptions occur in the workflows, usability, and functionality of the program. It increases the time to record information and reduces the time of direct care to the patient [4,5,7,8,10-18].

When the right technology is successfully implemented, it can increase efficiency, decrease workloads, and provide time to perform direct care [19-21].

To improve these barriers, authors have proposed developing portable systems so that EHRs can be completed at the person's point of care [13,18].

EHRs at the point of care promote patient quality and safety [13,19,21-25]. There are portable systems like reading barcodes for administering medication or to identify patients within the application as well as systems with early warning scores. These features increase patient safety. In addition, this system reduces errors from data transcription and latency time (information is recorded at the time it is obtained), which results in more accurate records. Quality is improved, and relying on remembering information is avoided. It provides accurate, real-time information, which improves accuracy and therefore patient safety. Resources that do not make it easy to complete EHRs at the patient's bedside make it easier to make errors. For example, a health care professional cannot know if the biobiotic prophylaxis inserted by surgical intervention is correct if the electronic medical history is not registered [13,23-25].

Bedside patient EHRs also promote communication and continuity of care [8,14,22-26]. The possibility to access information within the room facilitates communication between professionals, with the patients, and with their families, making them part of their health process and thus increasing their satisfaction and continuity of care [8,14,23-27].

Furthermore, EHRs at the point of care improve workflows. They decrease travel time and avoid transcription, which save time as they reduce the time spent recording and increase the time spent with the patient [25-27].

However, authors have found that nurses themselves have varying perceptions and opinions about bedside patient EHRs. They are more satisfied with mobile devices and prefer to use them for complex patients (because they require a large amount of information) and to record vital signs and blood products [14,25].

However, they prefer not to use them in 2 specific situations: for noncomplicated patients, because their records are simple (require little data), and to admit the patient, because it takes a long time (it is better to find a quiet place to do so). Other reasons include that, while they are recording the information, they have to answer questions from patients and families (disrupting concentration), they feel that they are not giving good patient care (because they are concentrating on the screen),

and they feel that it does not offer them the opportunity to disconnect (the post-record gives them time and space for this). In addition, bibliography authors also argue that documentation is not a high priority in the nurses' activities, something that needs to be done straight away. Professionals do not feel that documentation affects the timeliness of patient care [17,24,27,28].

In other words, for nurses, bedside patient EHRs require mental and technical skills. However, they do not perceive that, if they do not complete them, it will impair the quality of the care provided.

The articles consulted refer to a variety of devices, and there is no consensus on the most appropriate resource for completing EHRs at the point of care. Evaluating this is complex and is related to many different factors in each center, such as the Wi-Fi connection, access, and identification system [22-26].

In the literature consulted, no scientific articles were found in Spanish states that evaluated the effectiveness of a practical experience with EHRs at the point of care.

From other information sources, there are 2 examples of practical experiences implemented in the Spanish state: the projects at the Hospital Infanta Cristina de Madrid and Osakidetza Hospital in the Basque Country [29,30].

All of these arguments show that it is important for nurses to decrease recording and technology time, as it is seen as a complementary, administrative, and bureaucratic task.

Nurses may perceive introducing technology as an increase in these bureaucratic tasks and a detriment to direct patient care. This situation may involve less physical contact with the patient and more time spent with electronic devices [14].

One of the nurse theories that can help understand this is Dr. Ray's Theory of Bureaucratic Care, which focuses on nursing in complex organizations, such as hospitals [31]. He explains that, if we rely only on administrative theories or on theories focused solely on the patient-nurse relationship, the organization will not be able to adapt to new needs. Economic benefits and competitiveness (bureaucracy) prevail in contemporary organizations. However, there has been a resurgence of nursing as an art of science focused on human care (patient-nurse relationship). The Bureaucratic Care Theory clarifies the meaning of human care in complex organizations, placing it at the center as it is an essential part of hospital management. Human care self-organizes, interrelates, and interconnects with each of its parts, placing spiritual-ethical care at the center (the engine that moves the nursing practice) and around it the bureaucratic factors, such as the educational, physical, sociocultural, legal, economic, political, and technological factors [31-33].

It is necessary to take into account all these arguments and reflections when implementing any changes to EHRs. We believe that completing EHRs at the point of care with tablets can meet nurses' needs and expectations. It can strike a balance between nurses' need to provide direct patient care and the requirement to complete EHRs, and thus improve and facilitate care.

General Objective

The aim of this study was to evaluate whether completing EHRs at the point of care with tablets reduces nurses' time spent on records compared with the current system. This time difference can be used for direct patient care.

Specific Objectives

The first specific objective was to describe the sociodemographic variables (age and gender) of the nursing professionals in internal medicine units participating in the study, depending on the shift and the initial and final perceptions of the project.

The second was to describe and evaluate the sociodemographic variables (age and gender), shift, exclusion criteria, and initial and final perceptions of nurses working in internal medicine units who were not able to participate in the study.

The third was to compare the difference between the age groups of participants and those not participating in the study.

The fourth was to assess whether the implementation of this new record system decreases the time spent on patients compared with the current system and depending on the shift.

The results of this study were used to inform the impact of tablet use on workflows and to detect improvements to facilitate patient bedside registration.

Methods

Overview

A before-after single time motion study was conducted for 3 months (February 2017 to April 2017) in the internal medicine units of a regional hospital. A time and motion study is a quantitative data collection method in which the observer records the time and actions and movements of the participants. This type of study is often used for computer applications [34,35].

A total of 310 hours of observation and 130 observations of complete patient records were performed. Each observation could last between 2 hours and 3 hours and were carried out in the afternoon hours of 3:00 PM and 5:30 PM and during the night, from 10:00 PM to 1:00 AM (at the beginning of the nurses' work shifts).

No sampling technique was applied. The sampling was of all the nurses working in the medical unit who met the inclusion criteria. A representative sample of all regional hospital nurses was ruled out because the working conditions did not allow a larger sample to be monitored. We therefore decided to focus on a single unit that would make research feasible in terms of time and resources. As the sample size was small, we decided to establish a minimum amount of observations for each participant in the control and experimental groups. The significance level was .05, and the beta risk was less than .2 in a bilateral contrast. The common standard deviation was 1.96, and therefore, the size would be 15 observations per control group and 15 observations for the experimental group to detect a difference ≥1.5 units, estimating a monitoring loss rate of

10%. These data were obtained with a sample size calculator [36].

Following these calculations, we decided to perform a minimum of 10 observations for each nurse, 5 for the control condition (current system) and 5 for the experimental condition (using the tablet). A total of 23 professionals could participate in this study for 3 months, as they were the only ones who had the minimum dedicated time to be able to be observed continuously; however, considering the inclusion criteria, there were 13 individuals who could participate satisfactorily. Each person could only be in 1 group.

The control group used a computer on wheels (current system), and the experimental group used the tablet.

The tools for gathering information were 2 ad hoc databases with the study variables.

Inclusion Criteria

The inclusion criteria were as follows. The nurse was hired and worked in a stable job. A minimum of 10 observations could be made within 3 months, as this would facilitate the adaptation and learning curve in the experimental phase. The nurse carried out care activities in internal medicine units. This is because it is a robust unit, in which the occupancy rate is more stable and the average hospital stay is greater than 3 days. The nurse works the afternoon or night shift; the night shifts are longer than the other shifts because the staffing and distribution of activities are different [37]. The nurse volunteered to participate in the study. The nurse had a certain level of competence according to Benner's skills acquisition model [38]: It is necessary that the nurse has worked more than 3 years in the hospital, because in this time, they can learn the functioning of the hospital and the EHRs as well as the protocols and procedures of the center [38]. The nurse can use information and communication technologies (ICT) [28].

Demographic and Descriptive Variables

The variables age, sex, perception, and work shift of the participants and nonparticipants in the study were analyzed. The participants were those who were able to use the tablets according to the inclusion criteria: a total of 13. The nonparticipants were those who did not use the tablets: a total of 10.

Perception was measured by asking the professionals, at the beginning and at the end of the project, what their perception was: positive, neutral, negative.

There were 2 options for the shift variable (afternoon and night), according to the most common work shift of that nurse.

The exclusion criteria (only for nonparticipants) were measured using 5 items (sick leave, change of service, experience <3 years, does not accept, cannot use ICT).

Principal Variables

The number of patients admitted was the total number of patients who were hospitalized at the time of observation. Although the medical unit has 14 beds available per nurse, only actual occupancy was measured.

The total time was the result of the sum of the "round" time and time to record the data in the EHR. A "round" time was defined as the routine established at the beginning of the afternoon and evening shifts when vital constants are taken, the nurse activities are standardized, and an overall assessment is made of the patients. There is no such routine in the morning.

The time spent per patient was obtained by dividing the total time by the number of patients admitted. In addition, it was identified as the main variable of the study, because it is more standardized.

Control Variables

The following exceptional situations were defined: exitus, vital emergency, hospital discharge. In these cases, observations were discarded.

Justification of Variables

The selection of variables was justified by a study conducted at a Toronto hospital that measured the time nurses spent completing EHRs. The variables studied were the total time, time spent per patient, and number of patients admitted. This study compared paper records versus electronic records [18].

Two studies have used the variables sex, age, nationality, and experience related to positive perception of EHRs [10,39].

A previous study ruled out observations in case of emergency or illness [22].

The rest of the variables were chosen to adapt the research to the study field, to the practical and real situation of the work units.

Data Analysis

Statistical analysis was carried out with SPSS 17.00 (IBM Corp, Armonk, NY), and a significance level of $P<.05$ was applied. For categorical variables, a frequency calculation and contingency table were performed using likelihood ratio analysis (n<60). For the quantitative variables, for each of the groups, the Shapiro-Wilk normality test (n<30) was performed. The mean, SD, and variance were analyzed using an analysis of variance (ANOVA), and 2-factor averages for samples (control and experimental groups) were contrasted using the Student *t* test (Wilcoxon test when the sample was not normal). Following this analysis, when the results showed differences in averages between groups, Cohen d was applied to measure the distance or effect between the groups [40].

Ethics Approval

This paper is part of the research for a doctoral thesis, with a favorable report from the Bellvitge Clinical Research Ethics Committee (reference PR191/16). To respect the privacy and confidentiality of the informants, the databases only worked with their coded names. Participants provided their verbal and written consent. Each participant was informed about the purpose of the study, the voluntariness to participate, and the right to leave at any time.

Results

Demographics and Qualitative Results

Of the 23 professionals stably working in the internal medicine unit, 13 (56%) were able to participate in the study, and 10 (43%) could not.

As we can see in Table 1, most nurses participating in the study were women (11/13, 84%) versus men (2/13, 15%). The average age was 38.08 (SD 1.40) years. The 2 groups were homogeneous in the variables studied; there were no statistically significant differences in age between sexes and shifts (sex Leven *P*=.04; Mann Whitney *P*=.37; shift Leven *P*=-.02; Mann Whitney *P*=.10).

Table 1. Demographics and qualitative results from participants and nonparticipants.

Variables	Participants (n=13), n (%)	Nonparticipants (n=10), n (%)
Sex		
Male	2 (15)	1 (10)
Female	11 (84)	9 (90)
Age (years)		
<26	0 (0)	2 (20)
26-30	1 (7)	1 (10)
31-35	3 (23)	0 (0)
36-40	4 (30)	1 (10)
41-45	5 (38)	2 (20)
46-50	0 (0)	1 (10)
>50	0 (0)	3 (30)
Initial perception		
Positive	6 (46)	6 (60)
Negative	4 (30)	3 (30)
Neutral	3 (23)	1 (10)
Final perception		
Positive	9 (69)	N/A[a]
Negative	1 (7)	N/A
Neutral	3 (23)	N/A
Shift		
Afternoon	7 (53)	5 (50)
Night	6 (46)	5 (50)
Exclusion criteria		
Work leave	N/A	2 (20)
Change from the unit	N/A	1 (10)
Experience <3 years	N/A	3 (30)
No accept	N/A	2 (20)
No ICT[b] basic level	N/A	2 (20)

[a]N/A: not applicable.

[b]ICT: information and communication technologies.

The initial and final perceptions about the project after participating in the project did not vary statistically by shift or age variables. However, in the group of only women (n=11), the results varied because the women's final perceptions improved after they participated in the study. An initial positive perception was present for 45% (5/11), and 63% (7/11) had a final positive perception, with significant differences (verisimilitude ratio *P*=.03) with a substantial Cramer coefficient (v=.65).

Also, most nonparticipating nurses were women (9/10, 90%) versus men (1/10, 10%). The average age was 42 (SD 4.3) years. Table 1 details the variables for initial perception, shift, and exclusion criteria.

If we look at the relationship of these variables with the initial perception toward the project, we could see that there is homogeneity with age and that the relationship with the exclusion criteria could not be established because the data were not robust. However, a statistically significant relationship was obtained with the shift variable (verisimilitude ratio *P*=.04), and 80% (4/5) of the nonparticipants in the afternoon shift had a positive perception of the project, which was better than the perception of the night shift (2/5, 40%); this difference had a very strong coefficient of association (Cramer v, *P*=.97). After the study had been explained, nonparticipating subjects in the night shift had a worse perception or acceptance.

As shown in Multimedia Appendix 1 and Multimedia Appendix 2, statistically significant differences in age were evident between the group of nonparticipants (excluding the group with experience <3 years) and the group with the afternoon shift and night shift participants. The ”nonparticipating“ average age was 49.57 (SD 2.92) years compared with the average ages of 37.71 (SD 2.33) years for the “afternoon shift participants” and 38.50 (SD 1.60) years for the “night shift participants” (Tukey afternoon *P*=.007; Tukey night *P*=.01).

Principal Findings

The quantitative variables were used to measure and compare the times required to carry out the “round” and EHRs between the control group (current system) and the experimental group (tablet).

Of the total sample, the mean total time obtained for the control group was 55.44 (SD 2.11) minutes, and there was an average 11.77 (SD 0.25) patients admitted. For the tablet group, the average total time was 48.30 (SD 2.24) minutes, and there was an average 11.37 (SD 0.28) patients admitted.

Comparing the time spent per patient (the main variable of the study), it was evident that the average time spent per patient was lower with the tablet group mean 4.22, SD 0.14 minutes) than with the control group (mean 4.66, SD 0.12 minutes); there was a statistically significant difference (W=3.20; *P*=.001) and a low effect (d=.44) between groups (Table 2).

Table 2. Comparison of the average time spent per patient for the entire sample.

Variable	Control		Tablet		Comparisons			
	Mean (SD)	SW[a]	Mean (SD)	SW	Mean difference (SD)	W test[b]	*P* value	Cohen d
Time spent per patient (minutes)	4.66 (0.12)	.07	4.22 (0.14)	.006	0.44 (0.13)	–3.20	.001	.44

[a]Shapiro Wilk *P*<.05.

[b]Wilcoxon.

However, if we focused on analyzing these variables while taking into account the shift, the results brought a nuance or specificity to these more general data. The mean number of patients admitted was homogeneous and similar in the control (12.63, SD 0.22 patients) and tablet (12.50, SD 0.39 patients) groups. There were no statistically significant differences in their distribution (W afternoon *P*=.22; W night *P*=.96).

The average total afternoon shift times were 44.83 (SD 2.21) minutes for the control group and 35.48 (SD 1.17) minutes for the tablet group. The control group's night shift lasted 67.83 (SD 2.22) minutes, and the tablet group’s night shift lasted 63.27 (SD 2.78) minutes. The comparison of the average factor for related samples showed that there were significant differences in the afternoon shift (t_{34}= 4.07, *P*<.001), with a high effect between groups (d=.93), but not in the night shift (t_{29}=1.29, *P*=.20).

In the afternoon shift for the control group, the average times spent per patient were 4.07 (SD 0.13) minutes in the control group and 3.47 (SD 0.10) minutes in the tablet group. In the night shift, the control group spent an average 5.36 (SD 0.13) minutes, and the tablet group spent an average 5.09 (SD 0.19) minutes. Comparison of the average factor for related samples showed that, in the afternoon shift, the mean time spent per patient was lower with the tablet group, with a statistically significant difference (t_{34}=3.82, *P*=.01), and there was a high effect (d=.77) between groups. However, the same results were not obtained in the night shift (t_{29}=1.16, *P*=.25; Table 3 and Table 4).

Table 3. Average comparisons of total time, number of patients admitted, and time spent per patient for the afternoon shift.

Variable	Control		Tablet		Average comparison			
	Mean (SD)	SW[a]	Mean (SD)	SW	Mean difference (SD)	Test	*P* value	Cohen d
Total time (minutes)	44.83 (2.21)	.055	35.48 (1.17)	.47	9.37 (2.30)	4.07[b]	.001	.93
Number of patients admitted	11.03 (0.38)	.03	10.40 (0.33)	.24	.62 (0.49)	–1.20[c]	.22	—[d]
Time spent per patient (minutes)	4.07 (.13)	.21	3.47 (.10)	.25	.60 (0.15)	3.82[b]	.01	.77

[a]Shapiro Wilk *P*<.05.

[b]Student *t* test.

[c]Wilcoxon.

[d]Not calculated.

Table 4. Average comparisons of total time, number of patients admitted, and time spent per patient for the night shift.

Variable	Control		Tablet		Average comparison		
	Mean (SD)	SW[a]	Mean (SD)	SW	Mean difference (SD)	Test	*P* value
Total time (minutes)	67.83 (2.22)	.38	63.27 (2.78)	.85	4.56 (3.52)	1.29[b]	.20
Number of patients admitted	12.63 (0.22)	.01	12.50 (0.39)	.00	0.13 (0.49)	–.04[c]	.96
Time spent per patient (minutes)	5.36 (0.13)	.16	5.09 (0.19)	.37	0.26 (0.22)	1.16[b]	.25

[a]Shapiro Wilk *P*<.05.

[b]Student *t*.

[c]Wilcoxon

Discussion

Principal Findings

The results confirm the overall objective of this study: Completing EHRs at the bedside with tablets reduces nurses' time spent recording compared with the current system in which nurses use a computer on wheels. This difference is due to improved workflows, as bedside EHR completion with a tablet avoids unnecessary travel, facilitates access to information, and avoids duplicating the work involved in data transcription. The same results were obtained for the afternoon shift but not for the night shift. Therefore, this registration system can meet the expectations of nurses and produce a positive impact on work dynamics since it covers 2 important needs for these professionals: saving time on bureaucratic tasks and having more time for care.

However, the perceptions of the participants and nonparticipants in the study did not always coincide with this premise. There was no initial broadly positive perception of this registration system. The nonparticipants were older than the participants, and of these, nonparticipants on the night shift had a worse perception toward the project.

Comparison With Prior Work

The literature consulted shows conflicting results regarding whether electronic records reduce the time spent on records by citing numerous benefits and limiting agents. However, there is consensus on the concept for time: For nurses, time is an important and present concept. During their working day, they carry out numerous activities, always keeping in mind the way of organizing these so they can do them all in their work shift. Nursing is a pragmatic profession, in which an activity has to have a certain result. However, it is also a profession involving contact and a relationship with the patient through providing care, which is the main axis and motivation of their profession. Therefore, saving time in bureaucratic activities, to have more time for human care, is a constant concern [2,4,7-15,33,34,37].

There are other similar studies that investigated the impact of EHRs in workflows. They argue that, even though EHRs have been implemented, paper and subsequent transcription of information are still used. There is unanimity in saying that these practices are not advisable but differ in the problems they can cause, such as increasing latency and transcription errors [18,19] and duplication of work process [41].

Other authors who have published articles related to EHRs at the point of care have reached similar conclusions. It takes less time to enter the records because there is no need to move to a different place to enter the record, and work is halved because there is no need to enter handwritten records in the computer [20,22-24].

It is difficult to make comparisons between this study and other similar studies in which the time spent on EHRs at the point of care was calculated and demonstrated that technology can reduce time in registration or administrative tasks by nurses [19,20]

due to the use of different research methods, variables, and resources. In the first article [19], the authors found a 30% reduction in time spent on records. The second [20] article obtained an increase of 6% in time devoted to direct care and a reduction of 12% time in administrative tasks. In this investigation, time saved using the tablets was 0.44 (SD 0.13) minutes; for the afternoon shift, it was 0.60 (SD 0.15) minutes. In the investigation by Wong et al [19], differences between hospitalization units and the night shift were not considered due to the observations occurring from 9 AM to 5 PM. In our investigation, we found differences between shifts. In 2 other studies [19,20], they compared paper versus EHRs at the point of care, and in our study, we compared EHRs that were not performed next to the patient versus EHRs at the point of care.

In the literature that we consulted, 3 studies studied the relationship between sociodemographic variables and the impact and acceptance by nurses of information systems. However, they were studied to achieve different objectives and with different results, but similar conclusions could be reached. One study found consistent data related to previous experience in the use of computers with more favorable attitudes toward EHRs [10]. Another investigation revealed that performance expectancy and social influence were significant predictors of nursing information system usage intentions and suggested that nursing managers should promote usage [39]. The third research study obtained more negative results and claimed that the use of mobile devices intensifies the negative effects of usability problems related to EHRs, and they suggested different actions related to improving the usability and interface of the applications. Moreover, they referred to the relationship that the nurse's experience has with pressure and distress [1]. According to our results, the initial perception of the nurses was not unanimously positive, and the findings related a worst perception with age and the night shift. We do not know if age may be related to inexperience in the use of ICT, the interface, or fear of change or the unknown. Moreover, it was not possible to determine the factors related to the shift.

In order to improve acceptance, we agree with other authors [1,3,10,39], that nursing management should promote bedside EHRs and explain their benefits but should also offer continuing education courses and sufficient training in information systems to all nurses. Resources should be invested to improve the interface through the participation of nurses in its development.

Study Limitations

The time the researchers dedicated to the study was limited because it was carried out outside of working hours, there were few resources available to carry out the research (one tablet), and there were huge difficulties in making 10 observations with the same participant. For these reasons, there are limitations in terms of sample size and nonprobabilistic sampling type. To ensure external validity, this research could be repeated with a larger sample and random sampling in the future.

However, no greater control was taken over the confusion variables that could cause the results to vary. This would improve internal validity.

Finally, the number of nursing activities recorded per patient (eg, vital constants, pain, catheter, oxygen therapy, health education, scales) was not quantified at the time of observations. This was to avoid the effect of the observer and to promote informal acceptance of the study by the participants. The cost benefit of implementing these measures would have to be assessed for future studies.

Conclusions

This investigation allows us to know the impact EHRs at the point of care can have on workflows.

First, our findings determined that, with bedside EHRs, the time spent in registration by the nursing staff decreases, because of reduced movements and elimination of data transcription. Because EHR completion at the bedside eliminates unnecessary work that does not add value, care is improved. So, we think EHRs at the point of care should be the future or natural method for nursing to undertake.

On the other hand, our study explored sociodemographic and qualitative variables associated with this new registration system. It allows us to identify the factors that can make people reluctant to participate in a technological project, and we performed actions aimed at solving them in order to anticipate the possible obstacles. Otherwise, we could make statements that are wrong or biased and do not correspond with reality nor solve the problem. However, more studies with a larger sample would be needed to improve the validity of these results. It is proposed that training in the different work platforms and the participation of nurses are fundamental axes that any institution should consider.

This research is the result of the preparation of a doctoral thesis, and these findings will be triangulated in the second part when a qualitative phenomenological study is conducted on the experience and perceptions of nurses with EHRs at the point of care.

Taking into account that the subject of this research is quite unknown, especially in the Spanish territory; daily use of technology is part of our society; and nurses are the capital of hospital work templates, it is imperative to go in-depth in similar studies to provide more information and allow us to develop systems that promote patient care.

Acknowledgments

The data used in this article were obtained for the thesis entitled “Evaluation of implanting a point-of-care record system with tablets in a local hospital” directed by Ms Lina Cristina Casadó-Marín, PhD, and Mr José Oriol Romaní-Alfonso, PhD. We are also grateful to Mr Arnau Besora (Pere Virgili Fundation) and Mr Francesc Valls, PhD (Universitat Rovira i Virgili), for their support with the statistical analysis. We would like to thank the Universitat Rovira i Virgili language service and Mr Chris Place

(friend and freelance translator) for his English revision of the text. Finally, we would like to acknowledge the collaboration of the Nursing Department and Research Manager (Mr Alejandro Rodríguez-Molinero, PhD) and especially the nurses of the internal medicine unit of Consorci Hospitalari Alt Penedès-Garraf.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Average analysis of variance (ANOVA) comparison of the age variable.
[DOC File , 29 KB - humanfactors_v9i1e30512_app1.doc]

Multimedia Appendix 2
Multiple comparisons of averages between the groups for the age variable.
[DOC File , 29 KB - humanfactors_v9i1e30512_app2.doc]

References

1. Heponiemi T, Kaihlanen A, Gluschkoff K, Saranto K, Nissinen S, Laukka E, et al. The association between using a mobile version of an electronic health record and the well-being of nurses: cross-sectional survey study. JMIR Med Inform 2021 Jul 06;9(7):e28729 [FREE Full text] [doi: 10.2196/28729] [Medline: 34255704]
2. Rojas CL, Seckman CA. The informatics nurse specialist role in electronic health record usability evaluation. Comput Inform Nurs 2014 May;32(5):214-220. [doi: 10.1097/CIN.0000000000000042] [Medline: 24473121]
3. Zadvinskis IM, Garvey Smith J, Yen P. Nurses' experience with health information technology: longitudinal qualitative study. JMIR Med Inform 2018 Jun 26;6(2):e38 [FREE Full text] [doi: 10.2196/medinform.8734] [Medline: 29945862]
4. Gregory D, Buckner M. Point-of-care technology: integration for improved delivery of care. Crit Care Nurs Q 2014;37(3):268-272. [doi: 10.1097/CNQ.0000000000000030] [Medline: 24896558]
5. Kent B, Redley B, Wickramasinghe N, Nguyen L, Taylor NJ, Moghimi H, et al. Exploring nurses' reactions to a novel technology to support acute health care delivery. J Clin Nurs 2015 Aug;24(15-16):2340-2351. [doi: 10.1111/jocn.12881] [Medline: 26041122]
6. Elgin KH, Bergero C. Technology and the bedside nurse: an exploration and review of implications for practice. Nurs Clin North Am 2015 Jun;50(2):227-239. [doi: 10.1016/j.cnur.2015.02.001] [Medline: 25999067]
7. Domingos CS, Boscarol GT, Brinati LM, Dias AC, Souza CC, Salgado PDO. A aplicação do processo de enfermagem informatizado: revisão integrativa. eglobal 2017 Oct 01;16(4):603. [doi: 10.6018/eglobal.16.4.278061]
8. Smallheer BA. Technology and monitoring patients at the bedside. Nurs Clin North Am 2015 Jun;50(2):257-268. [doi: 10.1016/j.cnur.2015.02.004] [Medline: 25999069]
9. Sousa PAFD, Sasso GTMD, Barra DCC. Contribuições dos registros eletrônicos para a segurança do paciente em terapia intensiva: uma revisão integrativa. Texto contexto - enferm 2012 Dec;21(4):971-979. [doi: 10.1590/s0104-07072012000400030]
10. Kelley TF, Brandon DH, Docherty SL. Electronic nursing documentation as a strategy to improve quality of patient care. J Nurs Scholarsh 2011 Jun;43(2):154-162. [doi: 10.1111/j.1547-5069.2011.01397.x] [Medline: 21605319]
11. Dykes PC, Carroll DL, Benoit A, Coakley A, Chang F, Empoliti J, et al. A randomized trial of standardized nursing patient assessment using wireless devices. AMIA Annu Symp Proc 2007 Oct 11:206-210 [FREE Full text] [Medline: 18693827]
12. Medeiros A, Rezende L, Santos S, Nascimento N, Costa T. Applicability of electronic medical records in health sector: an integrative review. Revista de Enfermagem UFPE on line 2014;8(11):3992-4000 [FREE Full text] [doi: 10.5205/1981-8963-v8i11a13624p3992-4000-2014]
13. Wager KA, Schaffner MJ, Foulois B, Swanson Kazley A, Parker C, Walo H. Comparison of the quality and timeliness of vital signs data using three different data-entry devices. Comput Inform Nurs 2010;28(4):205-212. [doi: 10.1097/NCN.0b013e3181e1df19] [Medline: 20571371]
14. Sockolow PS, Rogers M, Bowles KH, Hand KE, George J. Challenges and facilitators to nurse use of a guideline-based nursing information system: recommendations for nurse executives. Appl Nurs Res 2014 Feb;27(1):25-32. [doi: 10.1016/j.apnr.2013.10.005] [Medline: 24360777]
15. Cornell P, Riordan M, Herrin-Griffith D. Transforming nursing workflow, part 2: the impact of technology on nurse activities. J Nurs Adm 2010 Oct;40(10):432-439. [doi: 10.1097/NNA.0b013e3181f2eb3f] [Medline: 20859094]
16. Gocsik T, Barton AJ. Workflow optimization: essential in the electronic health record journey. Clin Nurse Spec 2014;28(3):132-134. [doi: 10.1097/NUR.0000000000000048] [Medline: 24714427]
17. Lee S, McElmurry B. Capturing nursing care workflow disruptions: comparison between nursing and physician workflows. Comput Inform Nurs 2010;28(3):151-9; quiz 160. [doi: 10.1097/NCN.0b013e3181d77d3e] [Medline: 20431357]

18. Yeung MS, Lapinsky SE, Granton JT, Doran DM, Cafazzo JA. Examining nursing vital signs documentation workflow: barriers and opportunities in general internal medicine units. J Clin Nurs 2012 Apr;21(7-8):975-982. [doi: 10.1111/j.1365-2702.2011.03937.x] [Medline: 22243491]
19. Wong D, Bonnici T, Knight J, Gerry S, Turton J, Watkinson P. A ward-based time study of paper and electronic documentation for recording vital sign observations. J Am Med Inform Assoc 2017 Jul 01;24(4):717-721 [FREE Full text] [doi: 10.1093/jamia/ocw186] [Medline: 28339626]
20. Banner L, Olney CM. Automated clinical documentation: does it allow nurses more time for patient care? Comput Inform Nurs 2009;27(2):75-81. [doi: 10.1097/NCN.0b013e318197287d] [Medline: 21685832]
21. Lang A, Simmonds M, Pinchin J, Sharples S, Dunn L, Clarke S, et al. The impact of an electronic patient bedside observation and handover system on clinical practice: mixed-methods evaluation. JMIR Med Inform 2019 Mar 06;7(1):e11678 [FREE Full text] [doi: 10.2196/11678] [Medline: 30839278]
22. Fleischmann R, Duhm J, Hupperts H, Brandt SA. Tablet computers with mobile electronic medical records enhance clinical routine and promote bedside time: a controlled prospective crossover study. J Neurol 2015 Mar;262(3):532-540 [FREE Full text] [doi: 10.1007/s00415-014-7581-7] [Medline: 25476692]
23. Thompson BW. The transforming effect of handheld computers on nursing practice. Nurs Adm Q 2005;29(4):308-314. [doi: 10.1097/00006216-200510000-00004] [Medline: 16260994]
24. Kohle-Ersher A, Chatterjee P, Osmanbeyoglu HU, Hochheiser H, Bartos C. Evaluating the barriers to point-of-care documentation for nursing staff. Comput Inform Nurs 2012 Mar;30(3):126-133. [doi: 10.1097/NCN.0b013e3182343f14] [Medline: 22024972]
25. Stuttle B. Computers in nursing: patient data at the point of care. Primary Health Care 2009 Jun 06;19(5):32-33. [doi: 10.7748/phc.19.5.32.s26]
26. Carlson E, Catrambone C, Oder K, Nauseda S, Fogg L, Garcia B, et al. Point-of-care technology supports bedside documentation. J Nurs Adm 2010 Sep;40(9):360-365. [doi: 10.1097/NNA.0b013e3181ee4248] [Medline: 20798618]
27. Blake H. Mobile technology: streamlining practice and improving care. Br J Community Nurs 2013 Sep;18(9):430, 432. [doi: 10.12968/bjcn.2013.18.9.430] [Medline: 24005486]
28. Haller G, Haller DM, Courvoisier DS, Lovis C. Handheld vs. laptop computers for electronic data collection in clinical research: a crossover randomized trial. J Am Med Inform Assoc 2009;16(5):651-659 [FREE Full text] [doi: 10.1197/jamia.M3041] [Medline: 19567799]
29. El H. Infanta Cristina mejora la seguridad del paciente con el uso de tablets. APISCAM. 2016 Feb 17. URL: http://apiscam.blogspot.com/2016/02/gobcm-el-h-infanta-cristina-mejora-la.html [accessed 2021-08-06]
30. Ozakidetza, galardona con el Premio Nacional de Informática y Salud 2015. Euskadi. 2016. URL: http://www.euskadi.eus/gobierno-vasco/contenidos/noticia/2016_01_21_30444/es_30444/30444.html [accessed 2021-08-06]
31. Ray MA. The theory of bureaucratic caring for nursing practice in the organizational culture. Nurs Adm Q 1989;13(2):31-42. [doi: 10.1097/00006216-198901320-00007] [Medline: 2922119]
32. Coffman S. Teoría de la atención burocrática de Marilyn Anne Ray. In: Alligood MR, Marriner-Tomey A, editors. Modelos y teorias en enfermeria, 7th edition. Barcelona, Spain: Elsevier; 2011.
33. Turkel MC. Continuing education: Dr. Marilyn Ray's Theory of Bureaucratic Caring. Int J Hum Caring 2007 Jun 01;11(4):57-75. [doi: 10.20467/1091-5710.11.4.57]
34. Baker KM, Magee MF, Smith KM. Understanding nursing workflow for inpatient education delivery: time and motion study. JMIR Nurs 2019;2(1):e15658 [FREE Full text] [doi: 10.2196/15658] [Medline: 34345775]
35. Lopetegui M, Yen P, Lai A, Jeffries J, Embi P, Payne P. Time motion studies in healthcare: what are we talking about? J Biomed Inform 2014 Jun;49:292-299 [FREE Full text] [doi: 10.1016/j.jbi.2014.02.017] [Medline: 24607863]
36. Marrugat J, Vila J. Calculadora de tamaño muestral GRANMO. Antaviana. 2012. URL: https://www.imim.cat/ofertadeserveis/software-public/granmo/ [accessed 2021-08-06]
37. Pérez-Martí M, Casadó Marín LC, Romaní Alfonso JO. Estudio situacional de los registros electrónicos en una unidad de medicina interna: barreras y oportunidades. eglobal 2018 Jun 29;17(3):180. [doi: 10.6018/eglobal.17.3.284211]
38. Benner P. From Novice to Expert: Excellence and Power in Clinical Nursing Practice. Reading, MA: Addison-Wesley; 1984.
39. Chang I, Lin P, Chen T, Chang C. Cultural impact on the intention to use nursing information systems of nurses in Taiwan and China: survey and analysis. J Med Internet Res 2020 Aug 12;22(8):e18078 [FREE Full text] [doi: 10.2196/18078] [Medline: 32784174]
40. Manso Martíez ME, Sáchez Lóez MDP, Flores IC. Salud y sobrecarga percibida en personas cuidadoras familiares de una zona rural. Clínica y Salud 2013 Mar;24(1):37-45. [doi: 10.5093/cl2013a5]
41. Baumann LA, Baker J, Elshaug AG. The impact of electronic health record systems on clinical documentation times: A systematic review. Health Policy 2018 Aug;122(8):827-836. [doi: 10.1016/j.healthpol.2018.05.014] [Medline: 29895467]

Abbreviations

ANOVA: analysis of variance

EHR: electronic health record
ICT: information and communication technologies

Edited by A Kushniruk; submitted 18.05.21; peer-reviewed by R Ciorap, T Ntalindwa; comments to author 10.07.21; revised version received 26.08.21; accepted 08.10.21; published 10.02.22.

Please cite as:
Pérez-Martí M, Casadó-Marín L, Guillén-Villar A
Electronic Records With Tablets at the Point of Care in an Internal Medicine Unit: Before-After Time Motion Study
JMIR Hum Factors 2022;9(1):e30512
URL: https://humanfactors.jmir.org/2022/1/e30512
doi:10.2196/30512
PMID:35142624

Original Paper

Effects of Behavioral Economics–Based Messaging on Appointment Scheduling Through Patient Portals and Appointment Completion: Observational Study

Su-Ying Liang[1], PhD; Cheryl D Stults[1], PhD; Veena G Jones[2], MD; Qiwen Huang[1], MS; Jeremy Sutton[2]; Guy Tennyson[2], MBA; Albert S Chan[2,3], MD, MS

[1]Sutter Health Center for Health Systems Research, Palo Alto, CA, United States

[2]Clinical Leadership Team, Sutter Health, Sacramento, CA, United States

[3]Center for Biomedical Informatics Research, Stanford University, Stanford, CA, United States

Corresponding Author:
Su-Ying Liang, PhD
Sutter Health Center for Health Systems Research
795 El Camino Real, Ames Building
Palo Alto, CA,
United States
Phone: 1 650 691 6432
Email: liangs2@sutterhealth.org

Abstract

Background: Behavioral economics–based techniques have been an increasingly utilized method in health care to influence behavior change by modifying language in patient communication (through choice architecture and the framing of words). Patient portals are a key tool for facilitating patient engagement in their health, and interventions deployed via patient portals have been effective in improving utilization of preventive health services.

Objective: We examined the impacts of behavioral economics–based nudge health maintenance reminders on appointment scheduling through a patient portal and appointment completion for 2 preventive services: Medicare wellness visits and Pap smear.

Methods: We conducted a retrospective observational study using electronic health record data from an integrated health care system in Northern California. Nudge health maintenance reminders with behavioral economics–based language were implemented for all sites in November 2017 for Medicare wellness visits and for selected sites in February 2018 for Pap smears. We analyzed 125,369 health maintenance reminders for Medicare wellness visits, and 585,358 health maintenance reminders for Pap smear sent between January 2017 and February 2020. The primary outcomes were rate of appointments scheduled through the patient portal and appointment completion rate. We compared the outcomes between those who received the new, behavioral economics–based health maintenance reminders (the nudge group) and those who received the original, standard health maintenance reminders (the control group). We used segmented regression with interrupted time series to assess the immediate and gradual effect of the nudge for Medicare wellness visits, and we used logistic regression to assess the association of nudge health maintenance reminders, adjusting for the propensity to receive a nudge health maintenance reminder, for Pap smear.

Results: The rates of appointments scheduled through the patient portal were higher for nudge health maintenance reminder recipients than those for control health maintenance reminder recipients (Medicare wellness visits—nudge: 12,537/96,839, 13.0%; control: 2,769/28,530, 9.7%, *P*<.001; Pap smear—nudge: 8,239/287,149, 2.9%; control: 1,868/120,047, 1.6%; *P*<.001). Rates of appointment completion were higher for nudge health maintenance reminders for Pap smear (nudge: 67,399/287,149, 23.5% control: 20,393/120,047, 17.0%; *P*<.001) but were comparable for Medicare wellness visits (nudge: 49,835/96,839, 51.5% control: 14,781/28,530, 51.8%; *P*=.30). There was a marginally gradual effect of nudge on number of appointments scheduled through the patient portal for the overall Medicare wellness visits sample (at a monthly rate of 0.26%, *P*=.09), and a significant gradual effect among scheduled appointments (at a monthly rate of 0.46%, *P*=.04). For Pap smear, nudge health maintenance reminders were positively associated with number of appointments scheduled through the patient portal (overall sample: propensity adjusted odds ratio [OR] 1.62; 95% CI 1.50-1.74; among scheduled appointments: propensity adjusted OR 1.61, 95% CI 1.47-1.76) and with appointment completion (propensity adjusted OR 1.07; 1.04-1.10).

Conclusions: Nudges, a behavioral economics–based approach to providing health maintenance reminders, increased the number of appointments scheduled through the patient portal for Medicare wellness visits and Pap smear. Our study demonstrates that a simple approach—framing and modifying language in an electronic message—can have a significant and long-term impact on patient engagement and access to care.

(JMIR Hum Factors 2022;9(1):e34090) doi:10.2196/34090

KEYWORDS

access to care; behavioral economics; online; web-based appointment scheduling; health service; behavior; health care

Introduction

Health care has become more directly accessible to the patient, in part due to the Health Information Technology for Economic and Clinical Health Act of 2009 [1], and also, as a result of increasing consumer demands [2].When first introduced, patient portals provided patients with limited access to their medical records and adoption was low [3]. In part to meet meaningful use requirements of the Health Information Technology for Economic and Clinical Health Act, web-based electronic health record portal capabilities have expanded to appointment scheduling, displaying lab results or viewing encounter notes written by their physician [4], allowing bills payments, and facilitating communication with care teams, costs estimates for ambulatory services [5], and access to family records. Patient access rates to web-based patient portals have increased to 90% in some organizations [6]. Yet, access alone can only go so far in engaging patients with patient portals. Message construction and delivery also influence patient utilization of patient portals [7].

In the field of behavioral economics, researchers recognize that humans do not always act logically [8] in making choices and provide tools to help influence desired behaviors. One strategy includes choice architecture interventions, in which the presentation of options are altered to improve decision-making without restricting choice [9]. Subtle design changes, such as reordering choices, limiting options, or modifying a default can significantly influence behavior. When applying these tools and others, choice architects operate with a key tenet in mind: reduce friction in decision-making to make the desired path the one of least resistance [8,10]. In addition to the choice environment, behavioral economists also know that the framing of words can nudge individuals in a given direction of action [9]. For instance, messages that use common language [11], harness aversion to loss and regret [11,12], or embed social elements and devices [11-13] into their core are likely to be very effective at driving behavior change [11-13].

Over the past decade, these tenets of behavioral economics have been increasingly utilized in health care for both patients (cancer screenings [14], hospital appointment no-show rates [7], medication adherence for behavioral health [15], HIV testing rates [16], obesity and binge eating [17]) and clinicians (prescribing behavior [18,19], overtreatment of diabetes [20]), with 83 publications in the top 3 highest impact general medicine journals (Journal of the American Medical Association, The Lancet, The New England Journal of Medicine) from 1998 to 2018 [21]. A systematic review [22] found patient portal interventions to be effective in improving a few psychological outcomes, medication adherence, and preventive service use. Yet, to the best of our knowledge, behavioral economics has not been applied to health maintenance reminders sent through web-based patient portals to improve patient engagement and increase preventive service use for annual Medicare wellness visits and Pap smear. General use of the Medicare wellness visits has increased over time [23], with 7% [24] to 8.1% [25] of Medicare beneficiaries receiving an annual wellness visit in 2011 increasing to 16% in 2014 [24] and 23% in 2016 [25]. Although rates for Pap smear are much higher, with 83.7% of women age 21 to 65 years reporting having one within the past 3 years in 2018 [26], there is still room for improvement. Our objective was to determine if small changes in the wording of health maintenance reminders could alter patient completion of these preventive health services.

Methods

Setting

Sutter Health is a large not-for-profit health care system serving more than 3 million people annually across 100 rural and urban communities in Northern California. Sutter Health was the first health care system in the United States to implement Epic System's MyChart patient portal (My Health Online) in 2001 [27,28]. As of 2020, there have been over 3 million patients enrolled in My Health Online, which can be accessed via the website or through the mobile app, to communicate with their care team, refill prescriptions, view lab results, pay bills, and schedule appointments (including video visits).

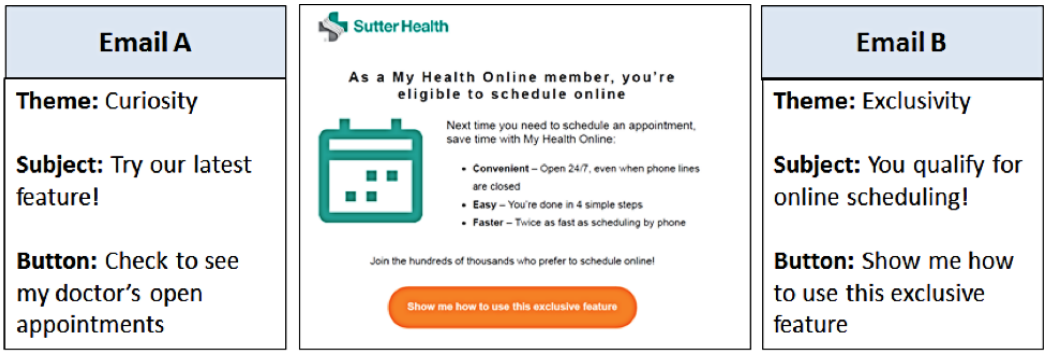
Pilot Testing: Behavioral Economics–Based Email Messaging

In September 2017, we conducted pilot testing to assess the application and effectiveness of behavioral economics–driven language in encouraging patients to schedule appointments using the patient portal with the 2 largest of the 5 Sutter Health medical foundations (Palo Alto Medical Foundation and Sutter Medical Foundation) in the Sacramento Valley. There were 775,000 Palo Alto Medical Foundation patients enrolled in My Health Online with 222,000 unique log-ins in September 2017 and 41,000 appointments scheduled through the patient portal. There were 436,000 Sutter Medical Foundation patients enrolled in My Health Online, with 103,000 unique log-ins in September 2017 and 13,000 appointments scheduled through the patient portal. We selected 2 sites to demonstrate that nudge health maintenance reminder could be useful for more than one patient population. This also allowed us to understand potential confounders and variables before scaling across the organization.

The pilot consisted of testing of 2 message options, followed by the roll-out of the message for which the most patients scheduled appointments through the patient portal. We tested 2 themes—curiosity (email A) and exclusivity (email B)—in patient messaging (Figure 1). We selected these themes based on guidance from behavioral economics experts at VAL Health based upon their previous experiences [29-31]. Among 550,000 My Health Online members from Palo Alto Medical Foundation and Sutter Medical Foundation who had never scheduled appointments through the patient portal, 3800 patients were randomized to receive the promotional emails. More patients, of both Sutter Medical Foundation and Palo Alto Medical Foundation, scheduled appointments through the patient portal after receiving email B, thus email B was selected for use in the next phase.

Email B was sent to the remaining Palo Alto Medical Foundation and Sutter Medical Foundation members who had never scheduled appointments through the portal (n=550,000) on November 10, 2017. We included a control group of 10% of Palo Alto Medical Foundation and Sutter Medical Foundation My Health Online users to see if there was any difference in rates of appointments scheduled through the patient portal compared to those for individuals who received nudge emails.

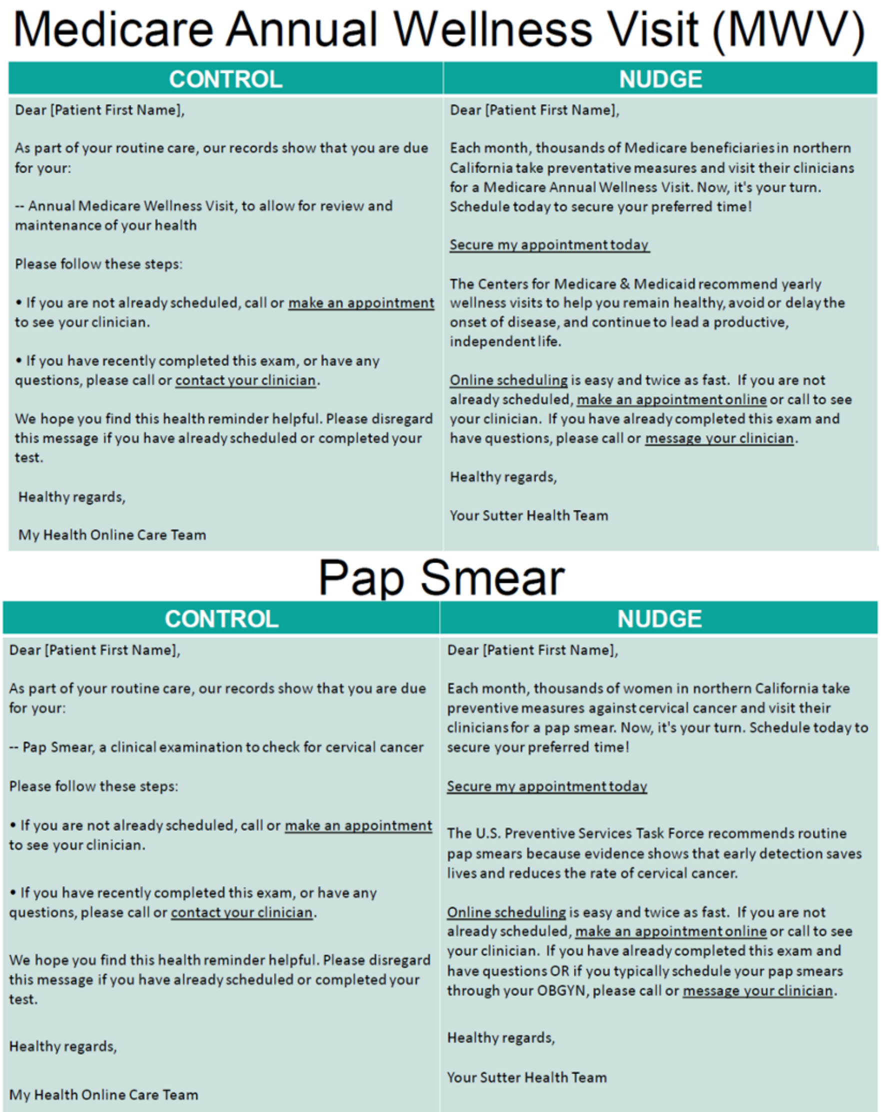
Figure 1. Themes of pilot behavioral economics–based email messages.



Application to Health Maintenance Reminders

The same behavioral economics concept was subsequently applied to health maintenance reminder messages in the patient portal. Nudge message wording was redesigned to use the exclusivity theme, with embedded functionality to click to schedule appointments through the patient portal for Medicare wellness visits and Pap smear (Figure 2). We selected Medicare wellness visits because it was already part of an active initiative at Sutter Health to increase utilization and we hoped that this type of message would better connect with patients who were eligible for Medicare wellness visits. We selected Pap smear so that we could examine how scheduling functions may impact screening completion and disease prevention. Nudge health maintenance reminders for Medicare wellness visits were implemented for all 5 Sutter medical foundations on November 15, 2017 (Figure 3). Nudge health maintenance reminders for Pap smear were launched on February 18, 2018; patients at Palo Alto Medical Foundation and Sutter Medical Foundation received nudge health maintenance reminders, while patients at the other 3 Sutter medical foundations continued to receive standard health maintenance reminders (control). Health maintenance reminders for Medicare wellness visits and Pap smear were discontinued during the COVID-19 pandemic; therefore, we analyzed utilization between January 2017 and February 2020.

Figure 2. Standard (control) and behavioral economics–based (nudge) health maintenance reminders.



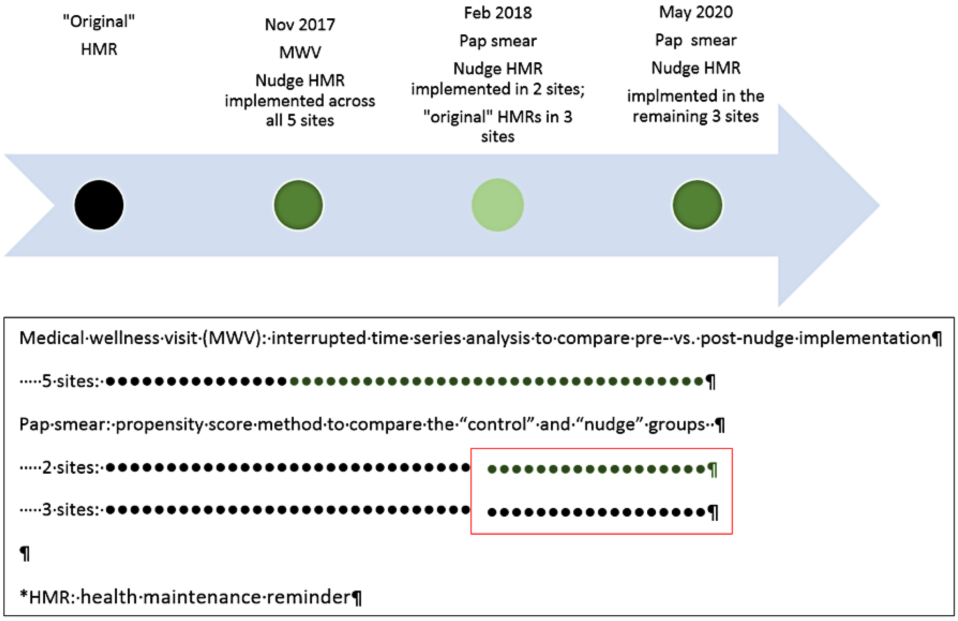
Medicare Annual Wellness Visit (MWV)

CONTROL	NUDGE
Dear [Patient First Name], As part of your routine care, our records show that you are due for your: -- Annual Medicare Wellness Visit, to allow for review and maintenance of your health Please follow these steps: • If you are not already scheduled, call or make an appointment to see your clinician. • If you have recently completed this exam, or have any questions, please call or contact your clinician. We hope you find this health reminder helpful. Please disregard this message if you have already scheduled or completed your test. Healthy regards, My Health Online Care Team	Dear [Patient First Name], Each month, thousands of Medicare beneficiaries in northern California take preventative measures and visit their clinicians for a Medicare Annual Wellness Visit. Now, it's your turn. Schedule today to secure your preferred time! Secure my appointment today The Centers for Medicare & Medicaid recommend yearly wellness visits to help you remain healthy, avoid or delay the onset of disease, and continue to lead a productive, independent life. Online scheduling is easy and twice as fast. If you are not already scheduled, make an appointment online or call to see your clinician. If you have already completed this exam and have questions, please call or message your clinician. Healthy regards, Your Sutter Health Team

Pap Smear

CONTROL	NUDGE
Dear [Patient First Name], As part of your routine care, our records show that you are due for your: -- Pap Smear, a clinical examination to check for cervical cancer Please follow these steps: • If you are not already scheduled, call or make an appointment to see your clinician. • If you have recently completed this exam, or have any questions, please call or contact your clinician. We hope you find this health reminder helpful. Please disregard this message if you have already scheduled or completed your test. Healthy regards, My Health Online Care Team	Dear [Patient First Name], Each month, thousands of women in northern California take preventive measures against cervical cancer and visit their clinicians for a pap smear. Now, it's your turn. Schedule today to secure your preferred time! Secure my appointment today The U.S. Preventive Services Task Force recommends routine pap smears because evidence shows that early detection saves lives and reduces the rate of cervical cancer. Online scheduling is easy and twice as fast. If you are not already scheduled, make an appointment online or call to see your clinician. If you have already completed this exam and have questions OR if you typically schedule your pap smears through your OBGYN, please call or message your clinician. Healthy regards, Your Sutter Health Team

Figure 3. Study design. HMR: health maintenance reminder; MWV: Medicare wellness visit.



Statistical Analysis

Study Sample

For this retrospective, observational study with electronic health record data, we identified 125,369 Medicare wellness visit health maintenance reminders sent to 43,889 unique patients who were 65 years and older between January 1, 2017 and February 28, 2020, and we identified 585,358 health maintenance reminders for Pap smear sent to 288,152 unique patients who were 21 years and older during the same study period.

Measures and Statistical Approach

The primary outcomes were rates of appointments scheduled through the patient portal and completed. The predictor of interest in all analyses was the receipt of nudge health maintenance reminder. The 2-tailed *t* test was used to examine differences in the unadjusted proportions of appointments scheduled through the patient portal and appointment completion between the nudge and control groups. We also conducted a subgroup analysis focusing on patterns of proportion of appointments scheduled through the patient portal among scheduled appointments.

For Medicare wellness visits, we used segmented regression with interrupted time series analysis to assess the immediate and gradual impact of the nudge intervention on study outcomes. The advantage of this analytic approach is its ability to distinguish the intervention effect from the secular trend (ie, a trend change that would have happened even in the absence of the intervention). The unit of analysis was the month. Study outcomes were measured in a given month: percentage of appointments scheduled through the patient portal (*number of health maintenance reminders leading to appointment scheduled online* / *total number of health maintenance reminders*) and percentage of appointment completion (*number of health maintenance reminders with appointment completed* / *total number of health maintenance reminders*).

For Pap smear, after initial launch of nudge health maintenance reminders, there remained a mix of nudge and control health maintenance reminders between February 2018 and February 2020. We used logistic regression models to assess the association of nudge health maintenance reminders with study outcomes, adjusting for the propensity to receive a nudge health maintenance reminder and accounting for clustering within the patients. The unit of analysis was the health maintenance reminder. The outcomes were use of the patient portal to schedule an appointment (yes or no) and completion of the appointment (yes or no). The propensity score method was chosen to control for potential selection bias and confounding by factors associated with the intervention and the study outcomes [32,33]. We estimated the propensity to receive a nudge health maintenance reminder as a function of patient demographic characteristics (age and race/ethnicity: White, Black, Hispanic, Asian, other), Charlson comorbidity score at baseline, health care utilization (number of health care encounters including in-person visit, video visits, My Health Online messages, and telephone calls) in the 12 months prior to baseline, insurance type (preferred provider organization or fee-for-service, health maintenance organization, Medicare, Medicaid, other), and their primary care physician's service location. Baseline was defined as the date of the first health maintenance reminder that the patient received during the analysis timeframe. Linear and categorical specifications of the propensity score were evaluated to ensure the robustness of the results. Analyses were conducted in R (version 4.0.4; The R Project) and Stata (version 16.1; StataCorp LLC).

Ethics

The study was reviewed by the Sutter Health Institutional Review Board (SHIRB) and approved as a quality improvement study (IORG0004135).

Results

Sample Characteristics

Of 125,369 health maintenance reminders sent for Medicare wellness visits, 77% (96,839) were nudge health maintenance reminders (ie, using behavioral economics–based language). Of 125,369 Medicare wellness visits health maintenance reminder portal messages sent, 60.2% (75,407) led to an appointment being scheduled, with 12.2% (15,342) scheduled through the patient portal, and 51.5% (64,616) appointments completed. Of 585,358 health maintenance reminders sent for Pap smear, 49.1% (287,149) were nudge health maintenance reminders. Of 585,358 Pap smear health maintenance reminder portal messages sent, 21.9% (128,329) led to appointment being scheduled, 2.3% (13,259) scheduled through the patient portal, and 21.9% (128,255) appointments completed (Table 1).

Of 43,889 patients (age: mean 75 years, SD 7.2) included in the Medicare wellness visit analysis, 60.7% (26,662/43,889) were women. Approximately two-thirds (33,837/43,889, 77.1%) were White, and 41% (18,175/43,889) had no comorbid conditions. The mean number of encounters in the 12 months prior to the baseline date was 22 (SD 20.6). Of 288,152 women (age: mean 41 years, SD 12.8) included in the Pap smear analysis, there were diverse racial/ethnic groups (White: 141,426/288,152, 49.1%; Black: 8919/288,152, 3.1%; Hispanic: 37,398/288,152, 13.0%; Asian: 59,225/288,152, 20.6%; other: 41,184/288,152, 14.3%); 89% (255,796/288,152) had no comorbid conditions. The mean number of encounters in the 12 months prior to baseline was 7 (SD 12.0); 45.9% (132,151/288,152) had coverage through preferred provider organization or fee-for-service plans, and 14.4% (41,618/288,152) had coverage through health maintenance organization plans (Table 2).

Table 1. Health maintenance reminders sent and outcomes.

Health maintenance reminder	Medicare wellness visits (n=125,369), n (%)	Pap smear (n=585,358), n (%)
Type		
Nudge	96,839 (77.2)	287,149 (49.1)
Control	28,530 (22.8)	298,209 (50.9)
Appointment scheduled		
No	49,962 (39.8)	457,029 (78.1)
Yes	75,407 (60.2)	128,329 (21.9)
Scheduled through patient portal		
Yes	15,342 (12.2)	13,259 (2.3)
No	110,027 (87.8)	572,099 (97.7)
Appointment completed		
Yes	64,616 (51.5)	128,255 (21.9)
No	60,753 (48.5)	457,103 (78.1)

Table 2. Sample characteristics.

Characteristic	Medicare wellness visits (n=43,889), n (%)	Pap smear (n=288,152), n (%)
Age (years), mean (SD)	75.3 (7.2)	40.9 (12.8)
Gender, n (%)		
Male	17,267 (39.3)	0 (0)
Female (%)	26,622 (60.7)	288,152 (100)
Race/ethnicity, n (%)		
White	33,837 (77.1)	141,426 (49.1)
Black	1121 (2.6)	8919 (3.1)
Hispanic	2947 (6.7)	37,398 (13.0)
Asian	3476 (7.9)	59,225 (20.6)
American Indian or Alaska Native/Pacific Islander or Native Hawaiian	198 (0.4)	1639 (0.6)
Multirace	347 (0.8)	5819 (2.0)
Unknown	1963 (4.5)	33,726 (11.7)
Comorbidity score at baseline, n (%)		
0	18,175 (41.4)	255,796 (88.8)
1	8462 (19.3)	22,167 (7.7)
2	7027 (16.0)	6882 (2.4)
3+	10,225 (23.3)	3307 (1.1)
Health care utilization at baseline, n (%)		
Encounters[a]	22.3 (20.6)	6.6 (12.0)
Insurance type, n (%)		
Health maintenance organization	373 (0.9)	41,618 (14.4)
Medicaid or Medi-Cal	36 (0.1)	9713 (3.4)
Medicare fee-for-service	28,291 (64.4)	7626 (2.7)
Medicare health maintenance organization	14,168 (32.3)	1045 (0.4)
Preferred provider organization or fee-for-service	750 (1.7)	132,151 (45.9)
Other/Unknown	271 (0.6)	95,999 (33.2)
Region (primary care physician's location at baseline), n (%)		
Region A	12,864 (29.3)	143,518 (49.8)
Region B	4133 (9.4)	19,969 (6.9)
Region C	24,402 (55.6)	62,988 (21.9)
Region D	1681 (3.8)	18,587 (6.5)
Region E	763 (1.7)	16,966 (5.8)
Other or no primary care physician	46 (0.1)	26,135 (9.1)

[a]Encounters included in-person visits, video visits, My Health Online message, and telephone.

Appointment Scheduling and Completion

For Medicare wellness visits, there was an increasing trend in proportion of appointments scheduled through the patient portal (Figure 4), and for Pap smear, appointments scheduled through the patient portal and appointment completion in the nudge group were consistently higher than those for the control group throughout the study period.

When comparing the unadjusted rates, we observed that a higher percentage of patients scheduled Medicare wellness visits through the patient portal after nudge implementation (nudge) than before implementation (control) for the overall sample (nudge: 12,573/96,839, 13.0%; control: 2769/28,530, 9.7%; $P<.001$) as well as for the subset with appointment scheduled (nudge: 12,573/58,371, 21.5%, control: 2769/17,036 16.3%; $P<.001$) (Table 3). A similar pattern for appointment scheduling through the patient portal was found for Pap smear between

February 2018 and February 2020 for those who received health maintenance reminders (nudge: 8239/287,149, 2.9%, control: 1868/120,047, 1.6%; *P*<.001) and for the subset with appointment scheduled (nudge: 12.2%, control: 9.2%, *P*<.001), and the rate of appointment completion for Pap smear was higher (*P*<.001) in the nudge group (67,399/287,149, 23.5%) than that in the control group (20,393/120,047, 17.0%), while rates of appointment completion for Medicare wellness visits were comparable (nudge: 49,835/96,839, 51.5%; control: 14,781/28,530, 51.8%; *P*=.30).

Figure 4. Trends of web-based scheduling and appointment completion from January 2017 to February 2020.

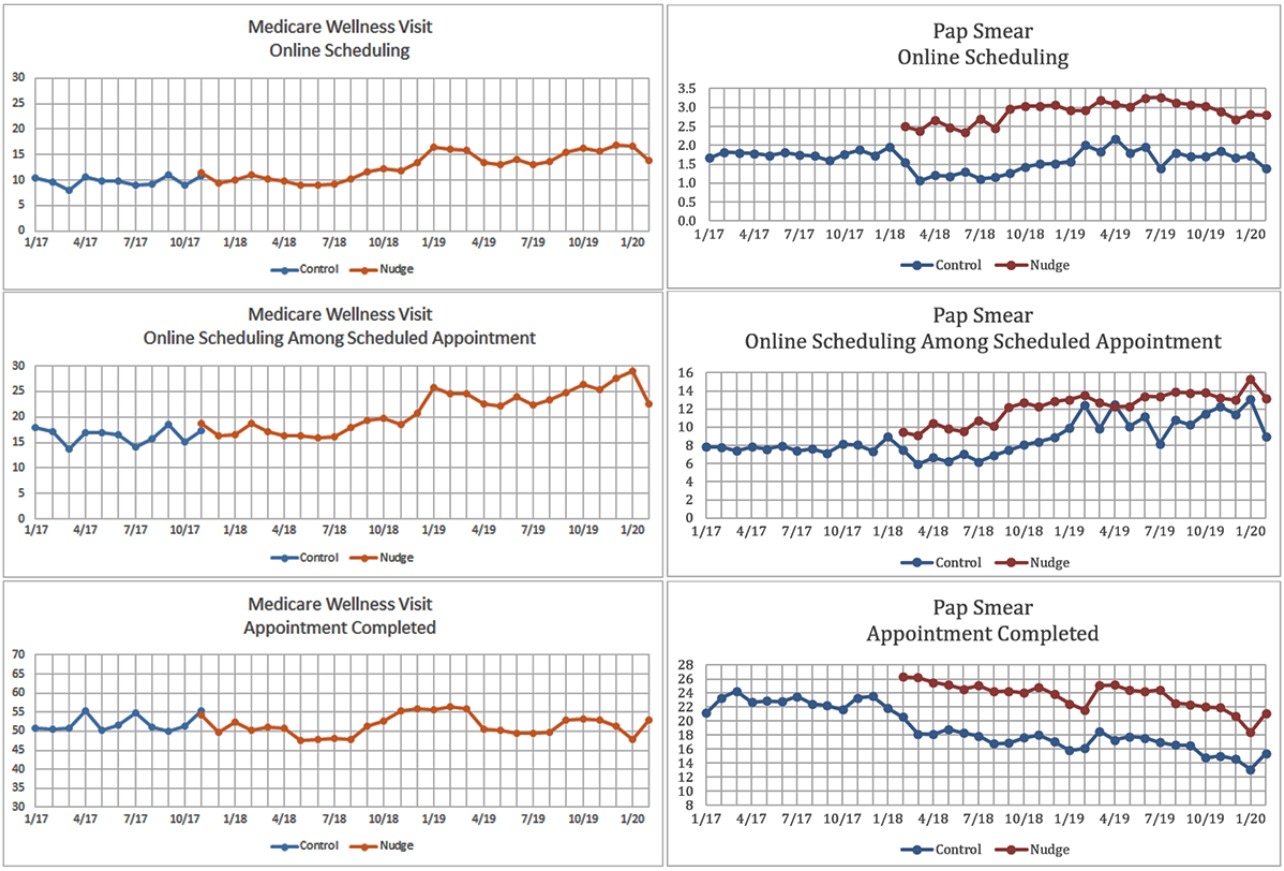


Table 3. Rates of appointment scheduling through the patient portal and appointment completion between nudge and control groups.

	Medicare wellness visits			Pap smear		
	Control (January 2017-November 2017)	Nudge (December 2017-February 2020)	*P* value	Control (February 2018-February 2020)	Nudge (February 2018-February 2020)	*P* value
Health maintenance reminders, n	28,530	96,839	—[a]	120,047	287,149	—
Appointments scheduled, n	17,036	58,371	—	20,402	67,445	—
Appointment scheduled through the patient portal, n	2769	12,573	—	1868	8239	—
Unadjusted[b], %	9.7	13.0	<.001	1.6	2.9	<.001
Adjusted[c], %	16.3	21.5	<.001	9.2	12.2	<.001
Appointments completed, n (%)	14,781 (51.8)	49,835 (51.8)	.30	20,393 (17.0)	67,399 (23.5)	<.001

[a]The comparison was not made.

[b]The percentage was calculated using the number of health maintenance reminders.

[c]The percentage was calculated using the number of appointments scheduled.

Nudge Effect

Findings from segmented regression for Medicare wellness visits (Table 4) suggested that there was a marginal gradual effect of nudge messaging on Medicare wellness visits scheduled through the patient portal (monthly rate 0.26%; *P*=.09). There was a statistically significant increase, at a rate of 0.46% per month (*P*=.04), in appointments scheduled through the patient portal. There was no immediate effect for either patterns in

appointments scheduled though the patient portal or for patterns in appointments completed for Medicare wellness visits.

The odds of scheduling a Pap smear for patients who received nudge health maintenance reminders were 1.62 times those for patients who received control health maintenance reminders (propensity adjusted odds ratio [OR] 1.62, 95% CI 1.50-1.74). A similar association was observed among the patients with scheduled appointments (propensity adjusted OR 1.61, 95% CI 1.47-1.76). Nudge health maintenance reminders were associated with a 7% increase in the odds of appointment completion for Pap smear (propensity adjusted OR 1.07, 95% CI 1.04-1.10).

We evaluated linear and categorical specifications (quintiles and deciles) of the propensity score. The categorical specification in deciles was selected, although sensitivity analyses using different specifications of the propensity score did not change our conclusions. Nudge health maintenance reminders were associated with a higher rate of scheduling through the patient portal for all health maintenance reminders and for patients who received health maintenance reminders and who scheduled appointments. For all health maintenance reminders, ORs ranged between 1.39-1.63 (quintile specification: adjusted OR 1.63, 95% CI 1.51-1.75; linear specification: adjusted OR 1.39, 95% CI 1.32-1.47). For patients with scheduled appointments, ORs ranged between 1.47-1.56 (quintile specification: adjusted OR 1.56, 95% CI 1.44-1.70; linear specification: adjusted OR 1.47, 95% CI 1.36-1.58). Nudge health maintenance reminders were associated with 10% to 17% increases in the odds of appointment completion for Pap smear (quintile specification: adjusted OR 1.17, 95% CI 1.14-1.20; linear specification: adjusted OR 1.10, 95% CI 1.07-1.12).

Table 4. Regression analysis results.

	Medicare wellness visits, coefficient (SE)				Pap smear[a], odds ratio (95% CI)	
	Preintervention level (Baseline)	Preintervention slope (secular trend, per month)	Change in intercept (immediate effect)	Change in slope (gradual effect, per month)	Unadjusted	Propensity adjusted
Scheduled through the patient portal	9.421 (0.941)	0.001 (0.151)	−0.700 (0.971)	0.264 (0.155)	1.87 (1.78 to 1.97)	1.62 (1.50 to 1.74)
Scheduled through the patient portal among all scheduled	16.343 (1.305)	−0.043 (0.210)	−1.071 (1.345)	0.462 (0.215)	1.38 (1.31 to 1.46)	1.61 (1.47 to 1.76)
Appointment completed	50.687 (1.728)	−0.007 (0.279)	−0.201 (1.782)	0.052 (0.285)	1.50 (1.47 to 1.53)	1.07 (1.04 to 1.10)

[a]The reference group in the model is patients who received a standard (control) health maintenance reminder. The propensity to receive a nudge health maintenance reminder was estimated as a function of patient age at baseline, race/ethnicity, Charlson comorbidity score at baseline, number of encounters at baseline, insurance type, and service location and was categorized in deciles in the adjusted model.

Discussion

Principal Findings

To the best of our knowledge, this is the first attempt to employ behavioral economics with electronic medical record portal–based health maintenance reminders to improve patient engagement and the utilization of preventive health services. Our findings suggest that our intervention influenced patient behavior. Simple modifications to verbiage and the framing of messages and adding embedded scheduling functionality had an impact on patients' actions. Although we did not find an immediate effect of nudge on scheduling appointments through the patient portal, we observed a sustained effect over time especially among those who scheduled an appointment. We believe this is because once patients learned how to schedule appointments through the patient portal, they used the portal to schedule appointments from that time onward. Similarly, we found that nudge health maintenance reminders for Pap smear were associated with a 65% to 68% increase in the odds of scheduling appointments through the patient portal, compared to control health maintenance reminders during the same study timeframe.

Appointments have typically been scheduled over the phone or in person, which can be resource intensive. Patient portals with web-based scheduling improves patient convenience, flexibility, and communications with providers, while reducing administrative burdens [34,35]. Furthermore, the COVID-19 pandemic has increased preference for and use of telehealth and features such as web-based scheduling. Behavioral economics–based health maintenance reminders are low-cost, effective, and operationally feasible. Once designed and pilot tested, they can be centrally deployed to all eligible patients with relatively low administrative burden and costs.

We acknowledge the very low rate for scheduling appointments through the patient portal for Pap smear (13,259/585,358, 2.3%) and posit that this may be due to a couple of factors. First, at Sutter Health, the majority of obstetric and gynecological practices do not allow appointments to be scheduled through the patient portal. Patients could schedule appointments with their primary care provider through the portal but could only type "Pap smear" in the free text box if they would like to receive one during that visit. Since there is no Pap smear visit scheduling type in the electronic medical record system, it added some complexity in retrieving these data—we identified the Pap smear procedure in subsequent visits. Nevertheless, even a small percentage means that more patients than those who would have with a standard message were able to receive this preventative screening.

Proactively engaging patients to schedule their appointments is the important first step to care management. Follow-through appointment completion is next. Our findings are mixed when examining the effect of nudge on appointment completion. We did not observe a significant effect on appointment completion for Medicare wellness visits (P=.30). The nudge effect for Pap smear was moderate, with a 7% increase, compared to control health maintenance reminders. The difference in nudge effects between Medicare wellness visits and Pap smear may be attributable to differences in patient populations to some degree. Exploratory stratified analyses suggested that patients who were younger (vs those 65 years and older) and Asian (vs other race or ethnicity groups) had the highest Pap smear appointment completion rate and most improvement from receiving nudge messages. Compared to the Medicare wellness visit sample, the Pap smear sample represented a younger patient population, with relatively more Asian individuals. We also explored if the number of health maintenance reminders sent plays a role in appointment completion. Among those who received multiple health maintenance reminders, subsequent health maintenance reminders were associated with higher rates of appointment completion than those for the first health maintenance reminders. These patterns were similar for both Medicare wellness visits and Pap smear. Scheduling an appointment or procedure does not automatically lead to completion, and thus, completion is a more complex process than scheduling and likely to be influenced by additional factors, which requires further investigation.

Limitations

Our study has several limitations. First, our analyses are based on data from a large health care system in Northern California. Our nudge intervention was designed for patients with access to internet and email which limits the generalizability of specific estimates of the nudge effect. Second, the propensity score covariate–adjusted method was selected because it mitigates confounding in observational studies such as ours. We chose this method over other widely used propensity score–matching methods. Due to the case-control ratio in our data (approximately 2.4:1 between February 2018 and February 2020), the propensity score–matching approach would have resulted in a loss of statistical power. The rollout of nudge health maintenance reminders was designed with operational goals to encourage patient uptake of web-based tools and to facilitate scheduling, which is expected to lead to better care management and health outcomes. As such, more cases than controls were enrolled. Different specifications of propensity score modeling may affect the results of our propensity score covariate–adjusted approach. We conducted sensitivity analyses using linear and nonlinear specifications of propensity score and our findings were consistent. Third, we focused on the overall patterns. Future research is needed to understand potential variability by specific subgroups (race/ethnicity, socioeconomic status) to inform targeted, culturally appropriate designs to maximize the benefit of nudges. The time of the year may have played a role in appointment scheduling and completion. Our study period limited the ability to examine the seasonal effect. We also recognize that the generalizability of these findings might be limited, with respect to application to other patient portals, as many have different functionalities and user experiences than those in the portal used in this study.

Conclusions

Nudges, a behavioral economics–based health maintenance reminder, improve web-based scheduling and subsequent appointment completion for Medicare wellness visits and Pap smear, with important long-term impacts. Given these results, Sutter Health implemented messages with behavioral economics–based language for all other health maintenance reminders on May 28, 2020. Future studies should explore why nudge worked for some patients and not others.

Acknowledgments

We are grateful to the contributions made by VAL Health, in study design, and to James Read and Kevin Hays for their assistance and expertise with electronic health record data extraction for this study.

Conflicts of Interest

None declared.

References

1. Health information technology for economic and clinical health (HITECH) act. US Department of Health and Human Services, and Ways and Means Committee. 2009. URL: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf [accessed 2022-03-10]
2. Patel V, Barker W, Siminerio E. rends in consumer access and use of electronic health information. Office of the National Coordinator for Health Information Technology. 2015. URL: https://www.healthit.gov/sites/default/files/briefs/oncdatabrief30_accesstrends_.pdf [accessed 2022-03-10]
3. Irizarry T, DeVito Dabbs A, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
4. Posada J, Potvin H, Cookson C. Opening up with Open Notes: writing notes in the era of full patient access. Acad Psychiatry 2021 Jul 26:1-5. [doi: 10.1007/s40596-021-01510-2] [Medline: 34309805]

5. Stults CD, Li J, Frosch DL, Krishnan H, Smith-McCurdy G, Jones VG, et al. Assessment of accuracy and usability of a fee estimator for ambulatory care in an integrated health care delivery network. JAMA Netw Open 2019 Dec 02;2(12):e1917445 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.17445] [Medline: 31834394]
6. Heath S. Patient portal adoption tops 90%, but strong patient use is needed. Patient Engagement HIT. URL: https://tinyurl.com/352tarjs [accessed 2020-04-29]
7. Berliner Senderey A, Kornitzer T, Lawrence G, Zysman H, Hallak Y, Ariely D, et al. It's how you say it: systematic A/B testing of digital messaging cut hospital no-show rates. PLoS One 2020 Jun 23;15(6):e0234817 [FREE Full text] [doi: 10.1371/journal.pone.0234817] [Medline: 32574181]
8. Volpp KG, Asch DA. Make the healthy choice the easy choice: using behavioral economics to advance a culture of health. QJM 2017 May 01;110(5):271-275 [FREE Full text] [doi: 10.1093/qjmed/hcw190] [Medline: 27803366]
9. Patel MS, Volpp KG, Asch DA. Nudge units to improve the delivery of health care. N Engl J Med 2018 Jan 18;378(3):214-216 [FREE Full text] [doi: 10.1056/NEJMp1712984] [Medline: 29342387]
10. Loewenstein G, Brennan T, Volpp KG. Asymmetric paternalism to improve health behaviors. JAMA 2007 Nov 28;298(20):2415-2417. [doi: 10.1001/jama.298.20.2415] [Medline: 18042920]
11. Asch D, Volpp K. Use behavioral economics to achieve wellness goals. Harvard Business Review. 2014. URL: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf [accessed 2022-03-10]
12. Emanuel EJ, Ubel PA, Kessler JB, Meyer G, Muller RW, Navathe AS, et al. Using behavioral economics to design physician incentives that deliver high-value care. Ann Intern Med 2016 Jan 19;164(2):114-119. [doi: 10.7326/M15-1330] [Medline: 26595370]
13. Asch DA, Rosin R. Engineering social incentives for health. N Engl J Med 2016 Dec 29;375(26):2511-2513. [doi: 10.1056/NEJMp1603978] [Medline: 28029924]
14. Purnell JQ, Thompson T, Kreuter MW, McBride TD. Behavioral economics: "nudging" underserved populations to be screened for cancer. Prev Chronic Dis 2015 Jan 15;12:E06 [FREE Full text] [doi: 10.5888/pcd12.140346] [Medline: 25590600]
15. Granek B, Evans A, Petit J, James MC, Ma Y, Loper M, et al. Feasibility of implementing a behavioral economics mobile health platform for individuals with behavioral health conditions. Health Technol 2021 Feb 25;11(3):505-510. [doi: 10.1007/s12553-020-00492-9]
16. MacCarthy S, Mendoza-Graf A, Wagner Z, L Barreras J, Kim A, Giguere R, et al. The acceptability and feasibility of a pilot study examining the impact of a mobile technology-based intervention informed by behavioral economics to improve HIV knowledge and testing frequency among Latinx sexual minority men and transgender women. BMC Public Health 2021 Feb 12;21(1):341 [FREE Full text] [doi: 10.1186/s12889-021-10335-5] [Medline: 33579242]
17. Graham AK, Munson SA, Reddy M, Neubert SW, Green EA, Chang A, et al. Integrating user-centered design and behavioral science to design a mobile intervention for obesity and binge eating: mixed methods analysis. JMIR Form Res 2021 May 10;5(5):e23809 [FREE Full text] [doi: 10.2196/23809] [Medline: 33970114]
18. Wang SY, Groene O. The effectiveness of behavioral economics-informed interventions on physician behavioral change: a systematic literature review. PLoS One 2020 Jun 4;15(6):e0234149 [FREE Full text] [doi: 10.1371/journal.pone.0234149] [Medline: 32497082]
19. Caturegli G, Materi J, Lombardo A, Milovanovic M, Yende N, Variava E, et al. Choice architecture-based prescribing tool for TB preventive therapy: a pilot study in South Africa. Public Health Action 2020 Sep 21;10(3):118-123 [FREE Full text] [doi: 10.5588/pha.20.0020] [Medline: 33134126]
20. Belli HM, Chokshi SK, Hegde R, Troxel AB, Blecker S, Testa PA, et al. Implementation of a behavioral economics electronic health Record (BE-EHR) module to reduce overtreatment of diabetes in older adults. J Gen Intern Med 2020 Nov 03;35(11):3254-3261 [FREE Full text] [doi: 10.1007/s11606-020-06119-z] [Medline: 32885374]
21. Reed KL, Harvey EM, Everly CJ. The intersection of behavioral economics and the general medicine literature. Am J Med 2021 Nov;134(11):1350-1356.e2 [FREE Full text] [doi: 10.1016/j.amjmed.2021.06.041] [Medline: 34343511]
22. Han H, Gleason KT, Sun C, Miller HN, Kang SJ, Chow S, et al. Using patient portals to improve patient outcomes: systematic review. JMIR Hum Factors 2019 Dec 19;6(4):e15038 [FREE Full text] [doi: 10.2196/15038] [Medline: 31855187]
23. Chung S, Lesser LI, Lauderdale DS, Johns NE, Palaniappan LP, Luft HS. Medicare annual preventive care visits: use increased among fee-for-service patients, but many do not participate. Health Aff (Millwood) 2015 Jan;34(1):11-20. [doi: 10.1377/hlthaff.2014.0483] [Medline: 25561639]
24. Ganguli I, Souza J, McWilliams JM, Mehrotra A. Trends in use of the US Medicare annual wellness visit, 2011-2014. JAMA 2017 Jun 06;317(21):2233-2235 [FREE Full text] [doi: 10.1001/jama.2017.4342] [Medline: 28423397]
25. Lind KE, Hildreth KL, Perraillon MC. Persistent disparities in medicare's annual wellness visit utilization. Med Care 2019 Dec;57(12):984-989. [doi: 10.1097/MLR.0000000000001229] [Medline: 31584462]
26. Bandi P, Minihan AK, Siegel RL, Islami F, Nargis N, Jemal A, et al. Updated review of major cancer risk factors and screening test use in the United States in 2018 and 2019, with a focus on smoking cessation. Cancer Epidemiol Biomarkers Prev 2021 May 19;30(7):1287-1299. [doi: 10.1158/1055-9965.epi-20-1754]
27. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006 Mar 01;13(2):121-126. [doi: 10.1197/jamia.m2025]

28. Tang PC, Black W, Buchanan J, Young CY, Hooper D, Lane SR, et al. PAMFOnline: integrating EHealth with an electronic medical record system. AMIA Annu Symp Proc 2003:644-648 [FREE Full text] [Medline: 14728253]
29. Behavioral economics: the last mile for health engagement. VALHealth. 2019. URL: https://tinyurl.com/2p9a397r [accessed 2021-12-14]
30. Achieving telemedicine adoption using behavioral economics. VALHealth. 2018. URL: https://tinyurl.com/2dk9j5bk [accessed 2021-12-15]
31. Increasing engagement with patient portals. VALHealth. 2018. URL: https://tinyurl.com/5n8e55yw [accessed 2021-12-04]
32. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011 May 31;46(3):399-424 [FREE Full text] [doi: 10.1080/00273171.2011.568786] [Medline: 21818162]
33. Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. JAMA 2007 Jan 17;297(3):278-285 [FREE Full text] [doi: 10.1001/jama.297.3.278] [Medline: 17227979]
34. Ganguli I, Orav EJ, Lupo C, Metlay JP, Sequist TD. Patient and visit characteristics associated with use of direct scheduling in primary care practices. JAMA Netw Open 2020 Aug 03;3(8):e209637 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.9637] [Medline: 32852551]
35. Zhao P, Yoo I, Lavoie J, Lavoie BJ, Simoes E. Web-based medical appointment systems: a systematic review. J Med Internet Res 2017 Apr 26;19(4):e134 [FREE Full text] [doi: 10.2196/jmir.6747] [Medline: 28446422]

Abbreviations

OR: odds ratio

Edited by A Kushniruk; submitted 06.10.21; peer-reviewed by R Krukowski, B Zhao; comments to author 24.10.21; revised version received 17.12.21; accepted 10.01.22; published 30.03.22.

Please cite as:
Liang SY, Stults CD, Jones VG, Huang Q, Sutton J, Tennyson G, Chan AS
Effects of Behavioral Economics–Based Messaging on Appointment Scheduling Through Patient Portals and Appointment Completion: Observational Study
JMIR Hum Factors 2022;9(1):e34090
URL: https://humanfactors.jmir.org/2022/1/e34090
doi:10.2196/34090
PMID:35353051

Original Paper

Personas for Better Targeted eHealth Technologies: User-Centered Design Approach

Iris ten Klooster[1], MSc; Jobke Wentzel[1,2,3], PhD; Floor Sieverink[4], PhD; Gerard Linssen[5], PhD; Robin Wesselink[6], MSc; Lisette van Gemert-Pijnen[1], PhD, Prof Dr

[1]Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands

[2]Department of Health and Social Studies, Windesheim University of Applied Sciences, Zwolle, Netherlands

[3]Research Group IT Innovations in Health Care, Windesheim University of Applied Sciences, Zwolle, Netherlands

[4]Carintreggeland, Hengelo, Netherlands

[5]Hospital Group Twente, Almelo and Hengelo, Netherlands

[6]Location AMC, Amsterdam University Medical Centers, Amsterdam, Netherlands

Corresponding Author:
Iris ten Klooster, MSc
Faculty of Behavioural, Management and Social Sciences
University of Twente
Drienerlolaan 5
Enschede, 7522 NB
Netherlands
Phone: 31 620730557
Email: i.tenklooster@utwente.nl

Abstract

Background: The full potential of eHealth technologies to support self-management and disease management for patients with chronic diseases is not being reached. A possible explanation for these lacking results is that during the development process, insufficient attention is paid to the needs, wishes, and context of the prospective end users. To overcome such issues, the user-centered design practice of creating personas is widely accepted to ensure the fit between a technology and the target group or end users throughout all phases of development.

Objective: In this study, we integrate several approaches to persona development into the Persona Approach Twente to attain a more holistic and structured approach that aligns with the iterative process of eHealth development.

Methods: In 3 steps, a secondary analysis was carried out on different parts of the data set using the Partitioning Around Medoids clustering method. First, we used health-related electronic patient record data only. Second, we added person-related data that were gathered through interviews and questionnaires. Third, we added log data.

Results: In the first step, 2 clusters were found, with average silhouette widths of 0.12 and 0.27. In the second step, again 2 clusters were found, with average silhouette widths of 0.08 and 0.12. In the third step, 3 clusters were identified, with average silhouette widths of 0.09, 0.12, and 0.04.

Conclusions: The Persona Approach Twente is applicable for mixed types of data and allows alignment of this user-centered design method to the iterative approach of eHealth development. A variety of characteristics can be used that stretches beyond (standardized) medical and demographic measurements. Challenges lie in data quality and fitness for (quantitative) clustering.

(JMIR Hum Factors 2022;9(1):e24172) doi:10.2196/24172

KEYWORDS

personas; clustering; heart failure; eHealth; user-centered design

Introduction

Although eHealth technologies are seen as an opportunity to support self-management and disease management for patients with chronic diseases, their actual use remains low [1]. As a result, the full potential of eHealth technologies is not being reached. A possible explanation for these lacking applications is that during the development process, insufficient attention is paid to the needs, wishes, and context of the prospective end users. To overcome such issues, user-centered design (UCD)

principles [2] provide tools to keep the intended user in the heart of the eHealth development process. The UCD practice of creating personas is widely accepted to ensure the fit between a technology and the target group or end users throughout all phases of development [3]. Personas represent fictive members of the target group and consist of a description of these potential users. By engaging with the personas, developers and project team members develop an eye for the characteristics of their target group [4]. One could say personas are a way to continuously communicate "who we are doing this for" to the team. In addition, the eHealth development team can, for example, anticipate on these personas to tailor educational messages [5] or to support adherence among several types of eHealth users [6]. The approaches that are described for creating personas are using 1 source of data, ignoring the variety and variability in data needed to create groups of end users that have similar characteristics.

Several frameworks advocate the use of multiple methods for data collection during the eHealth development process [7], for example, through interviews, questionnaires, and focus groups. Thus, mixed types of data from several sources are used during eHealth development, while persona creation often relies on limited data sources. First, the target group's health-related attributes form an important part of the personas in eHealth projects [8]: the risk of health complications, health-related activities that the prospective end users must undertake, the variation of symptoms in the target group, and tailoring options for medical treatment. These topics reflect factors that can be used to paint the "end user picture." Thus, health-related factors are the major contributors to the construction of eHealth personas. However, as research and experience in eHealth development progresses and matures, it has become obvious that an eHealth user should be characterized by more than just health status, and zooming in on health-related factors only tells a part of the users' story. Rather, a person who may be ill or has a chronic disease and aims to recover after surgery or disease or simply looks to preserve his/her health still has many more personal characteristics, likes, dislikes, or habits that are also relevant for understanding this person [9]. Therefore, second personas are created focusing on how a person wants, likes, or prefers to live life. LeRouge and colleagues [10] developed a conceptual model for identifying a broad range of user profiles and persona attributes from qualitative data. A related approach that considers characteristics beyond health factors is described by Vosbergen et al [5]. They have demonstrated how a variation in information needs can lead to personas (and consequently, technology design) that represent different ways in which people value and consume information. Similarly, there are many preferences, habits, and other variables beyond health/disease status and demographics that may be worthwhile to include in eHealth personas [11,12]. In these approaches, the personas result from a selection of relevant factors depending on more subjective experiences and tacit knowledge from experts. This can easily result in somewhat arbitrary decisions made on what to include in the persona. An approach that addresses this issue is proposed by Holden et al [13], using a quantitative cluster analysis on biopsychosocial survey data. In their approach, Holden et al [13] use qualitative data such as subjective eHealth literacy to describe the target group and distill personas that represent this group.

In addition to the use of health-related data and person-oriented data, we have noticed approaches in which server log data are used for identifying and describing user groups. Server log data are an automatic registration of, among others, the time, date, and activity that is carried out by the eHealth user within the system. An example is the identification of user groups based on activities within the eHealth system, resulting in personas characterized by activities that are most prominent within the clusters [14]. A more comprehensive approach is described in the study by Jones et al [15], in which activities within the system are expanded with information about the frequency, intensity, consistency, and demographics of the users. Using such data results in personas that include demographics of the users as well as users' engagement with an eHealth system. When this method is applied for identifying groups of eHealth users with chronic conditions, this approach itself can be expanded with log data related to monitored health values.

Overall, we see that there are several frameworks describing the steps in a very structured or less structured manner through which eHealth technologies can be developed. These frameworks are similar in that we see several data collection methods during the phases that are iteratively walked through to come to a technology that fits with the end users. In this sense, applying a framework in eHealth development and persona creation alike benefits from applying a broad lens to the user, technology, and context to ensure a good fit. The Center for eHealth Research (CeHRes) roadmap [7] describes such an approach, where research and development are guided through various design phases. This approach calls for holistic and value-driven development, focusing not only on the functionality and goal of eHealth technology but also accounting for users' motivations, abilities, circumstances, and context [7]. Personas fit well within this approach if we include relevant factors/characteristics for creating personas. However, the approaches for developing personas, as described above, only focus on 1 method for collecting data (eg, interviews, questionnaire data, log data), ignoring the variety of data collected during the UCD development processes. Therefore, we have studied how to develop a structured iterative approach for personas within the eHealth development process. Data from a previous study were used in which the phases of the CeHRes roadmap were completed, resulting in data that were collected through various methods (eg, interviews, questionnaires, log data).

Methods

Study Design

In this study, we have used a 3-step iterative approach to personas. In the first step, health-related data were used to develop the personas, using data from an electronic patient record (EPR). In the second step, these EPR data were enriched with person-related data that were gathered through interviews and questionnaires. In the third step, log data were added to the model to illustrate how personas can be further developed after log data are collected through a pilot study or after the eHealth

technology is launched and actually used by the end user. From now on, we refer to this iterative approach to eHealth development as the Persona Approach Twente (PAT). During this illustration of PAT, the focus is also on (1) how the approaches as described by Holden et al [13] and LeRouge et al [10] can be combined enabling the use of several data collection methods (quantitative and qualitative) for describing user groups and (2) the use of semiautomated methods for grouping the end users so that the arbitrary approach applied in previous studies for developing person-related personas is replaced by a more systematic approach. Thereby, we have aimed to contribute to achieving the full potential of eHealth technologies for chronic diseases.

Data Collection

Data collected in a previous study for the development of a telemonitoring application for people with heart failure were used, guided by the steps described in the CeHRes roadmap [7]. These data were gathered among 25 patients with mild to moderate chronic heart failure from the outpatient clinic of the Hospital Group Twente, Almelo and Hengelo, The Netherlands, of whom 13 were females (56%). Their mean age was 68 (SD 9) years, ranging between 46 and 82 years. Patients with a New York Heart Association (NYHA) functional classification 2 or 3 [16], with stable symptoms, and stable medication were included in this study. Persons admitted to the hospital within 1 month after data collection were excluded.

First, data from EPRs of the participants were used to collect health-related data such as NYHA classification and cerebrovascular accident or transient ischemic attack comorbidity. Second, quantitative data were collected through the 8-item eHealth Literacy Scale (eHEALS) questionnaire [17] to gain insight into the eHealth literacy status of the participants. Third, the 5-level 5-dimension Euro quality of life (EQ-5D-5L) questionnaire was used to gain insight into participants' quality of life, consisting of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [18]. Moreover, qualitative data regarding experiences in living with heart failure, technology use and trust, and motivation were collected through interviews with the participants. Based on these data, the iMediSense telemonitoring system (2016, Thales) was developed in another study [19]. A pilot study was conducted that was clinically supervised by cardiologists and nurse practitioners. In this pilot, patients were instructed to conduct measurements at least once daily for 60 days: diastolic blood pressure, systolic blood pressure, heart rate, and body weight. Further, they filled out an heart failure symptoms questionnaire. When measurements exceeded predefined ranges, alarms were generated. Nurse practitioners were instructed to view the generated alarms and react accordingly. The log data regarding the appointed symptoms, the alarms during the pilot study, and usage log data were used for the secondary analysis in this study. In Table 1, the aforementioned data collection methods are coupled with the variables that were collected through these methods. The variables also display the number of participants for whom a variable is known. Owing to the secondary analysis of this data set, not all variables were present or assessed among all participants.

Table 1. Data collection methods used in this study coupled with the variables that were collected through these methods and the number of participants for whom a variable is known.

Method, collected data	Variables (n)	Clustering
Electronic patient record		
Demographic	Gender (25), age (25)	Step 1 and 2 and 3
Medical	Cerebrovascular incident or transient ischemic attack comorbidity (25), chronic obstructive pulmonary disease comorbidity (25), diabetes comorbidity (25), left ventricular ejection fraction (25), heart failure with reduced ejection fraction (left ventricular ejection fraction <40%) (25), ischemic heart disease (25), hypertension (25), atrial fibrillation (25), New York Heart Association 2 or New York Heart Association 3 (25), heart failure hospitalization (25), cardiac resynchronization therapy defibrillator (25), estimated glomerular filtration rate (25), implantable cardioverter defibrillator (25)	Step 1 and 2 and 3
Interviews		
Technical	Smartphone ownership (23), personal computer ownership (22), tablet ownership (23), use of technology for entertainment (13), use of technology for social purposes (14), use of technology for gaining information (14)	Step 2 and 3
Demographic	Education type (7), children (13), grandchildren (5), divorce (13), marital status (16), employment (22)	Step 2 and 3
Health care specifics	Positive coping (25), negative coping (25), health-related goals (25), years ago diagnosed with heart failure (24)	Step 2 and 3
eHealth Literacy Scale questionnaire		
Capacity for engaging in eHealth	eHealth literacy (22)	Step 2 and 3
5-level 5-dimension Euro quality of life questionnaire		
Quality of life	Quality of life before using the telemonitoring technology (25)	Step 2 and 3
	Quality of life after using the telemonitoring technology (25)	Step 3
Log data of the pilot study		
Usage log data	Start new measurement (25), send symptoms measurement (25), send physical measurements (25), open history of measurements (25), contact care provider (25), open profile page (25), open user manual (25)	Step 3
Appointed symptoms	Restless, forgetful, and had a lacking concentration (25), a reduced effort level (25), a reduced appetite (25), a more than normal increase in fatigue (25), increased shortness of breath (25), cough or tickling cough (25), moisture in legs and abdominal distension (25), increased palpitations, fast paced heartbeat and chest pain (25)	Step 3
Generated alarms	Alarm for systolic blood pressure (24), alarms for diastolic blood pressure (24), alarms for heart rate (24), alarms for weight (24)	Step 3

Data Analysis

Before analyzing the data, the qualitative data collected through the semistructured interviews were coded by 2 independent coders (FS and JW) by using a combination of inductive and deductive coding [20]. First, the scheme of LeRouge et al [10] with codes related to technical, demographic, and health care specifics were used to code the interview data deductively. Subsequently, these codes were adapted and supplemented by means of inductive coding. After qualitative analysis, all resulting themes and variations were categorized into binary variables to enable cluster analysis. This means that if a theme consisted of several variations, multiple binary variables were created: 1 for every variation. For example, marital status was divided into 2 variables, namely, marriage (married or not married) and divorce (divorced or not divorced). Moreover, when a code was assigned to less than 5 quotes, then these were deleted from further analysis to reduce the influence of the missing values on the cluster results. Second, Shapiro-Wilk tests were performed to check whether variables were normally distributed [21]. We found that the variables age, capacity for engaging in eHealth, and estimated glomerular filtration rate were normally distributed (P>.05). The remaining variables were not normally distributed (P<.05) (Multimedia Appendix 1) and therefore log transformed before carrying out the cluster analyses.

Since data were both numerical and binary, distance matrices were created using Gower distances. Gower distances can handle these types of mixed data by using range-normalized Manhattan distances for quantitative data and Dice coefficient for nominal variables [22]. Subsequently, 3 cluster analyses were carried out using the Partitioning Around Medoids algorithm to develop personas related to 1 of the 3 steps in the PAT. A cluster analysis is a form of exploratory data analysis, where observations are divided into meaningful groups that share common

characteristics. The Partitioning Around Medoids algorithm was chosen since it fits with Gower distances, and the medoids can be used as "representatives" for the translation of clusters to personas. Medoids refer to observations that fall within a cluster for which the average dissimilarity between it and all the other members of the cluster is minimal. By using these representatives, we limit the influence of extreme values among the participants.

The analyses were conducted on 3 distinct parts of the same data set: (1) health-related data, (2) qualitative and quantitative health- and person-related data, and (3) qualitative and quantitative health- and person-related data, enriched with log data collected during the pilot study. All analyses were carried out using RStudio [23] and the R Cluster package [24], and results were visualized using the Ggplot2 package [25]. To estimate the optimal number of clusters, the average silhouette method was used. After conducting the cluster analyses, the medoids of the resulting clusters were used to describe personas. Table 1 summarizes which variables were included in the analysis for every step (1-3).

Ethics Approval

All participants gave permission for the use of these data and signed an informed consent form. Moreover, this study was ethically approved by the Behavioral, Management, and Social Sciences ethics committee (210111).

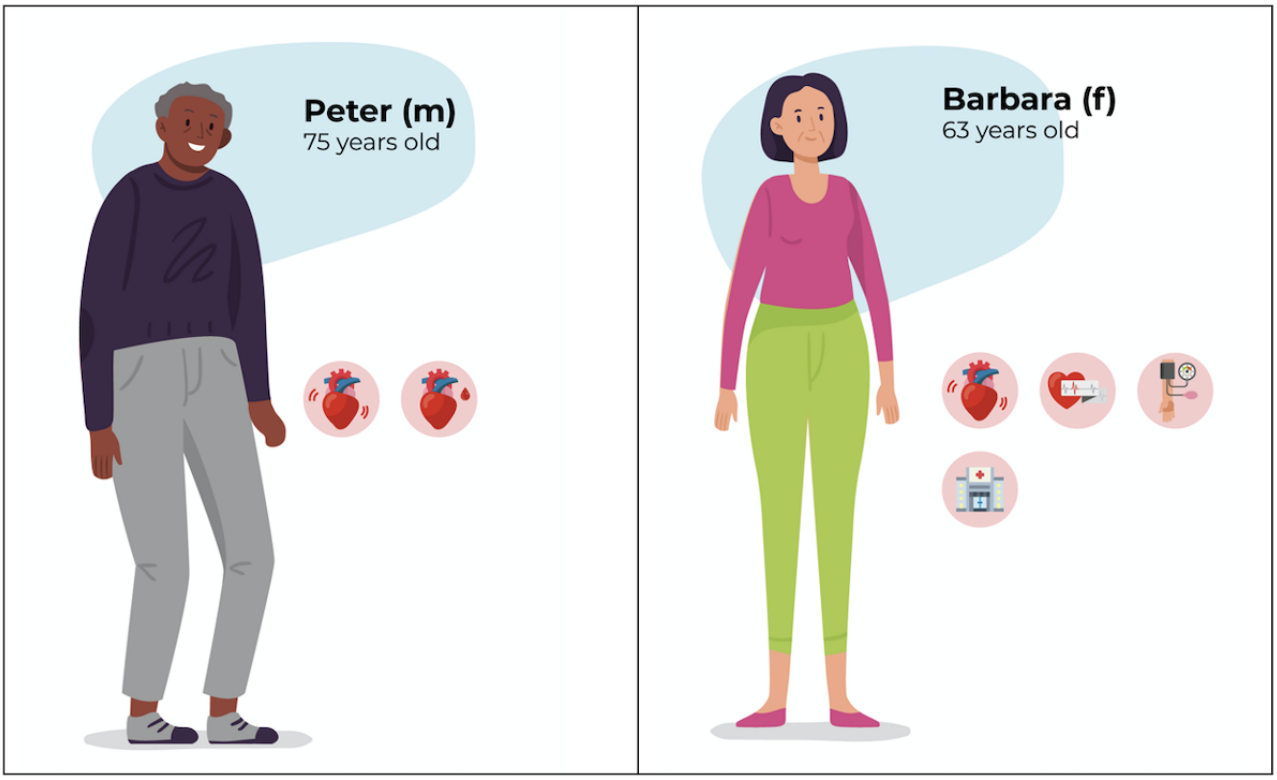
Results

Three cluster analyses were carried out that align with data collected through the (1) EPR (2) data enriched with interview and questionnaire data, and (3) the aforementioned data enriched with log data.

Clustering Health-Related Data

Figure S1 of Multimedia Appendix 2 shows the average silhouette widths for the number of clusters ranging from 2 to 10. Based on this figure, it was decided that the optimal number of clusters was 2, yielding an average silhouette width of 0.17. In total, 25 persons were divided into 2 clusters. The first cluster has an average silhouette width of 0.12 and consists of 17 persons, which is 68% (17/25) of the total number of persons. The second cluster has an average silhouette width of 0.27 and consists of 8 persons, which is 32% (8/25) of the total number of persons. The medoids of these clusters were used to translate these clusters in personas. Two personas were created using the variable values of these medoids, and these can be found in Figure 1 (the meaning of the symbols used in the persona descriptions are given in Multimedia Appendix 3). The first persona is Peter (representing cluster 1), who has heart failure with reduced ejection fraction and an ischemic etiology. Second, the persona Barbara represents cluster 2, who has heart failure with reduced ejection fraction, hypertension, atrial fibrillation, and her estimated glomerular filtration rate was reduced (43 mL/min/1.73 m^2). Barbara has had a prior hospitalization for heart failure.

Figure 1. Personas developed in the first step on the basis of clustering electronic patient record data. Multimedia Appendix 3 shows the meaning of the symbols used in the persona descriptions. The red background indicates the medical characteristics. f: female; m: male.



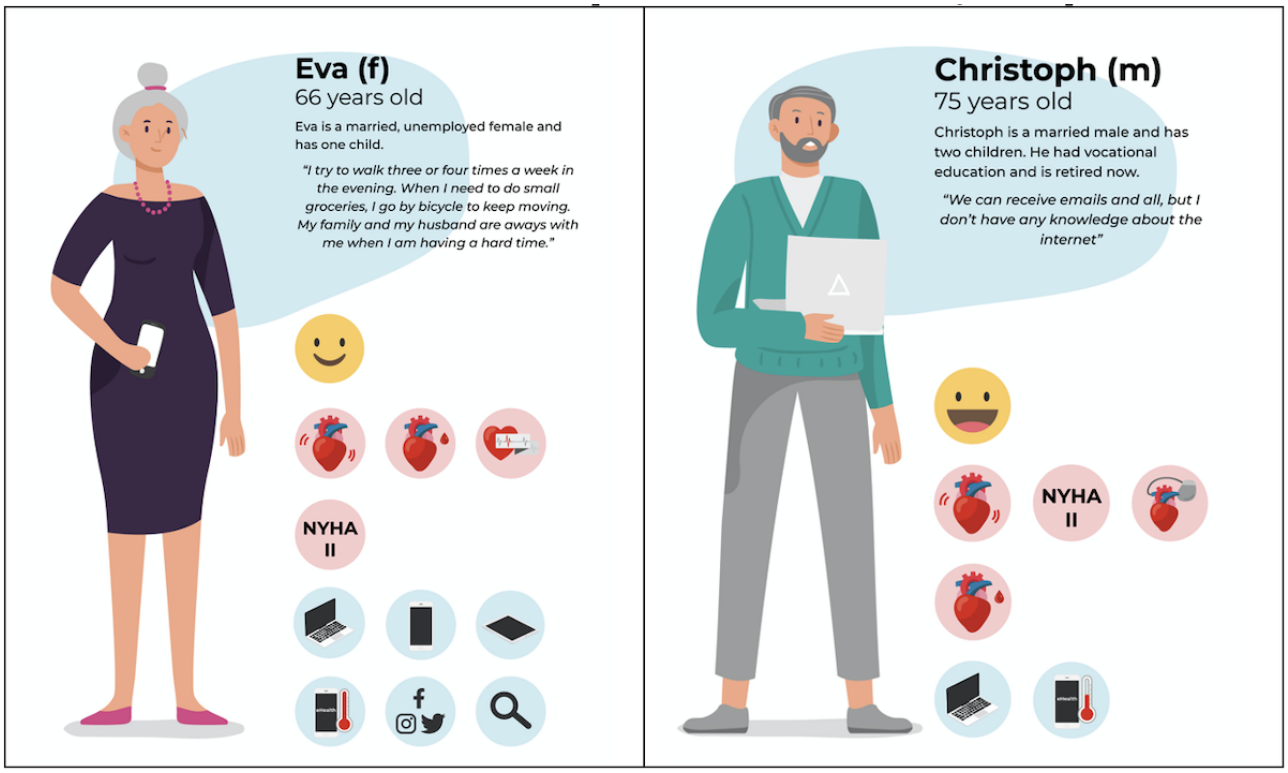
Clustering Health-Related Data Enriched With Person-Related Data

In the second step, we clustered the data set with health-related data, interview data, and the eHEALS questionnaire [17]. After the cluster analysis, an average silhouette plot was created yielding 2 clusters, and this can be found in Figure S2 of Multimedia Appendix 4. The corresponding average silhouette width for 2 clusters is 0.11.

Of the total of 25 persons, the first cluster consists of 10 persons (40%) with an average silhouette width of 0.08. The second cluster consists of 15 persons (60%) with an average silhouette width of 0.12. Persona descriptions were made based on the medoids within the 2 clusters, and these can be found in Figure 2. The first persona is Eva, who was diagnosed with heart failure with reduced ejection fraction of 33% and atrial fibrillation 2 years ago. Eva has a score of 10 on the EQ-5D-5L questionnaire on a scale from 5 to 25, indicating a good quality of life with slight problems or health issues. Eva mentioned 1 way of positive coping and 2 ways of negative coping. Eva owns a smartphone, computer, and a tablet. She uses this technology for social purposes (eg, social media) and for gaining information. Moreover, she has a mean score of 4 on the eHEALS questionnaire, indicating a moderately high capacity for engaging in eHealth. Correspondingly, Eva indicated that she has experience with eHealth technologies.

Christoph is a 75-year-old married male who had vocational education. He has 2 children and is currently unemployed. Christoph was diagnosed with heart failure with reduced ejection fraction 2 years ago. Besides, he has a left ventricular ejection fraction of 37% and an estimated glomerular filtration rate of 60 mL/min/1.73 m^2. Christoph has ischemic heart disease. Christoph has an implantable cardiac resynchronization therapy defibrillator or an implantable cardioverter defibrillator to support his heart function. Christoph has a score of 5 on the EQ-5D-5L questionnaire, indicating a good quality of life. He mentioned 2 ways of negative coping with problems. Christoph owns a computer but no smartphone or tablet. Moreover, he has a score 3 on the eHEALS questionnaire, indicating a moderate capacity for engaging in eHealth. Moreover, Christoph indicated that he has no skills in working with eHealth technologies.

Figure 2. Personas developed in the second step on the basis of clustering electronic patient record data, data from the interviews, the eHealth Literacy Scale questionnaire, and the 5-dimension 5-level Euro quality of life questionnaire. Multimedia Appendix 3 shows the meaning of the symbols used in the persona descriptions. The red background indicates the medical characteristics, and the blue background indicates the technical characteristics. f: female; m: male.



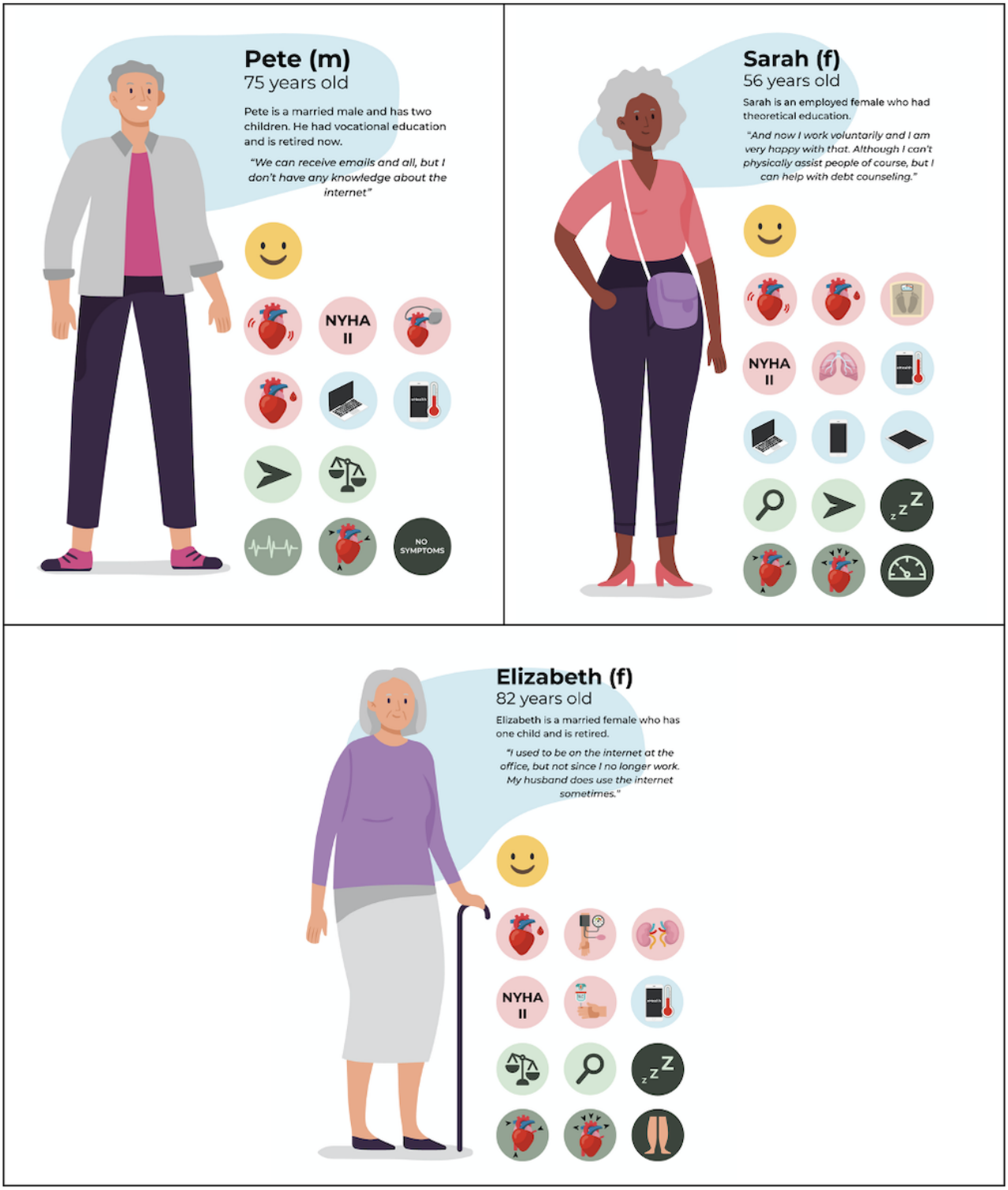
Clustering Health- and Person-Related Data Combined With Log Data

In the third step, we enriched the health- and person-related data with usage log data that are typically collected after the design phase. After the cluster analysis, an average silhouette plot was created yielding 3 clusters. This average silhouette plot can be found in Figure S3 in Multimedia Appendix 5. The corresponding average silhouette width for 3 clusters is 0.08. Of the 25 persons, the first cluster consists of 15 persons (60%) with an average silhouette width of 0.09. The second cluster consists of 5 persons (20%) with an average silhouette width of 0.12. The third cluster consists of 5 persons (20%) with an average silhouette width of 0.04. Persona descriptions were made based on the medoids within the 3 clusters, and these can be found in Figure 3.

The first persona is Pete (representing cluster 1) who was diagnosed with heart failure with a reduced ejection fraction

and an ischemic etiology 2 years ago. Pete did not mention any positive ways of coping and 2 ways of negative coping. Moreover, he has no smartphone or tablet, but he owns a computer. He had a score of 3 on the eHEALS questionnaire, indicating doubts in his skills to use information technology for health and mentioned that he has no skills in using eHealth technologies. During the pilot study, Pete indicated that he had no symptoms in the heart failure–symptoms questionnaire. Besides, alarms were mainly generated for heart rate (n=13) and diastolic blood pressure (n=10). During the pilot study, Pete showed a usage pattern in which only new measurements were started (n=77) and sent to the monitoring system (n=63). He visited his measurement history 1 time. Besides, he did not use other functionalities within iMediSense. His quality of life after using the monitoring system (EQ-5D-5L mean score of 5) did not change after using the monitoring technology.

Figure 3. Personas developed in the third step on the basis of clustering electronic patient record data, data from the interviews, the eHealth Literacy Scale questionnaire, the 5-level 5-dimension Euro quality of life questionnaire, and log data. Multimedia Appendix 3 shows the meaning of the symbols used in the persona descriptions. The red background indicates the medical characteristics, the blue background indicates that the technical characteristics, and the green background indicates the log data from the pilot of iMediSense. f: female; m: male; NYHA: New York Heart Association.



Sarah represents cluster 2, and she was diagnosed with heart failure and reduced ejection fraction 1 year ago. Her estimated glomerular filtration rate was 88 mL/min/1.73 m^2. Sarah has chronic obstructive pulmonary disease comorbidity and her goal is to maintain a stable weight. Sarah mentioned 1 way of coping positively and 3 ways of negative coping. Moreover, she owns a smartphone, tablet, and a computer. She finds her own skills in using of information technology for health reasonably high (eHEALS mean score of 4) and indicated that she has experience with eHealth technologies, but she does not see an added value. During the pilot study, Sarah indicated a mixed pattern of symptoms through the heart failure–symptoms questionnaire. She mentioned that she was restless, forgetful, and lacked concentration (n=4); she had reduced effort level (n=5), a

reduced appetite (n=4), a more than normal increase in fatigue (n=7), increased shortness of breath (n=3), and cough or tickling cough (n=2). Her quality of life increased slightly (EQ-5D-5L mean score of 13) compared to her quality of life before using iMediSense (EQ-5D-5L mean score of 12). During the pilot study, alarms were mainly generated for heartbeat (n=29). Besides, alarms for diastolic blood pressure were generated 13 times, and the alarms for systolic blood pressure were generated 17 times. In iMediSense, Sarah started a new measurement 52 times, sent the symptoms measurement 36 times, and the physical measurement 37 times. Besides, she opened her measurement history 54 times and opened her profile page 37 times. Furthermore, she visited other functionalities a few times.

The third persona is Elizabeth (representing cluster 3) who was diagnosed with heart failure 2 years ago. She has hypertension comorbid with diabetes. Moreover, she has an estimated glomerular filtration rate of 47 mL/min/1.73 m^2 and has been hospitalized before the current visit. Elizabeth has a score 3 on the eHEALS questionnaire, indicating doubts in her skills to use information technology for health. The main symptom that Elizabeth mentioned through the heart failure–symptoms questionnaire was moisture in legs and abdominal distension (n=37). During the pilot study, alarms were almost daily generated for systolic blood pressure (n=58) and diastolic blood pressure (n=43). In a much lower amount, alarms were generated for heart rate (n=5) and for weight (n=1). In iMediSense, Elizabeth shows a usage pattern in which she mainly started new measurements (n=165), sent the symptoms measurements (n=68), and looked into her measurement history (n=87). Her quality of life increased a little (EQ-5D-5L mean score of 8) compared to her quality of life before using iMediSense (EQ-5D-5L mean score of 7).

Discussion

Principal Findings

The practice of creating personas is widely accepted to ensure the fit between a technology and the target group or end users throughout all phases of development. Our demonstration of PAT shows that this approach can be used for developing personas through clustering mixed data in an iterative way to align with the process of eHealth development. This way, the richness of the persona increases as the development of an eHealth technology continues, while the use of a clustering algorithm partially ensures that these are objectively determined. PAT has the advantage that (1) the use of medoids makes the results easy to interpret, (2) mixed data can be used, and (3) personas can be iteratively developed. Below, we will elaborate on these advantages, and lastly, we will describe the disadvantages that we have encountered, along with a possible solution and direction for the future.

For the development of personas with heart failure by using PAT, we have used medoids as a method to find representatives for each group of users that have similar characteristics (clusters). These medoids have a minimum dissimilarity with other patients in the same cluster. Data from this representative patient (medoid) can be used for describing the persona. Holden et al [13] used comparative statistical tests between clusters to see on which variables these clusters differ. Subsequently, only the means of the variables that significantly differ are used to describe the personas [13]. PAT has several advantages compared to this approach. The first is that it is also suitable when the number of participants is low (which often occurs in the UCD process) since comparative statistical tests are also highly dependent on the number of participants. Second, PAT is less labor-intensive since it does not require to conduct comparative statistical tests. Third, it is easier to interpret, for example, a mean value of 0.5 on gender is difficult to interpret, which does not occur when medoids are used.

Besides the use of medoids, PAT allows including mixed data for persona development. In our demonstration of PAT on data collected in a project guided by the steps of the CeHRes roadmap, we were able to include data collected through questionnaires, interviews, EPR, and log data. This way, a more holistic understanding of the users can be reached. Moreover, including mixed data can be seen as an application of method triangulation [26]. For example, the NYHA classification of patients was extracted from EPRs, which is a description of the severity of the heart failure based on symptoms, and this classification ranges from I (no symptoms or limitations) to IV (severe limitations). However, the symptoms that a patient experiences and the way in which this limits the patient in his or her daily life might be understood in a more holistic way when adding data that are collected through another method.

Lastly, we saw how PAT aligned with the order in which data collection methods were deployed in the Twente Teach project. When applied during the development process, personas can be constantly updated based on newly collected data. This constantly updating of personas overlaps with the concept of Digital Twins [27]. The difference is that the current approach is focused on an up-to-date description of users on a group level, whereas Digital Twins are applied on an individual level. This ensures that the persona remains applicable and clear in the complex process of eHealth development. However, we do argue that the concept of “adaptive intelligence” should also be applied when PAT is used. This means that the personas are developed using an algorithm but that they become meaningful when domain knowledge is used for translating these personas into practical implications for targeting the users of the eHealth technology.

Although we found several advantages of PAT in this study, the results show that the quality of clusters decreases when qualitative data from the interviews are used in the cluster analysis (as expressed by the lower silhouette width). This, however, does not mean that the interview data are invaluable. Rather, it may imply that attention should be paid to what kinds of data are available or should be collected and how these are collected. Typically, health-related data are present for all patients included in a study, whereas the collection of more person-oriented characteristics of our patients or user groups is less standardized and defined. We argue that information about the person should be included, as health-related variables are measured more often and often in a more structured way, making them easier to use. The variables that stretch beyond health and tell us more about our user as a person, his/her background, circumstances, abilities, motivations, and values are at least as

valuable to measure and use to create personas. However, this study shows that data quality is an issue when modeling the personas, and this occurs more often in less standardized variables. This applies to many of the information types described in LeRouge et al's framework [10], which focuses on a broader context of eHealth user characteristics [10]. For example, technology use (technical specifics) or information-seeking attitudes (health care specifics) are potentially very relevant but are constructs that are rarely part of a standard and standardized medical assessment. To be able to use such possibly relevant variables, they should be measured in a more structured way.

Another possible remedy to this decreasing quality of clusters when adding qualitative data is to use domain knowledge for deciding which variables should be included in the cluster analysis or to summarize multiple variables into 1 variable (eg, use feature engineering or a factor analysis). However, since targeting eHealth users based on more than 1 variable is associated with a higher effectiveness of interventions [28], we state that a more systematic collection of person-oriented characteristics should be preferred. We argue that the steps below should be iteratively completed during the eHealth development process. These are also applicable in other contexts (eg, other target groups, when data are collected in a different order):

1. Collect data using a variety of methods and make sure that person-related variables are collected in a structured way.
2. Check whether variables are normally distributed and adjust analysis accordingly.
3. Carry out a cluster analysis to group participants into similar clusters.
4. Describe the clusters based on medoids and draw personas on the basis of the data that are known of these medoids.
5. Add qualitative data from these medoids to these personas to increase the richness of the persona descriptions.
6. Use domain knowledge to translate the personas into practical implications for the eHealth system to better target the eHealth to the users.

Limitations

Owing to the explorative design of this study, the small sample size of 1 clinical center, and the homogeneous sample accordingly, it remains unclear to what extent results can be generalized across patients with heart failure and other situations and groups of people. However, the focus of this study was to show how PAT might be used to develop personas; therefore, generalization was not a condition for useful results. Nevertheless, the question remains to what extent cluster results can still be used within a development process when collecting a larger amount of data from the group of end users. Moreover, usage log data of iMediSense could not be used because there was too little variation in that data: adherence was high (almost 100%) and the ways in which users could navigate through the platform were limited. It would be relevant to explore to what extent clustering results are of predictive value for the ways in which users navigate through a system, when indeed adherence and navigation patterns vary. Further, application of remote coaching and education to promote self-management may alter the clustering and predictive value of navigation through the system, which warrants further research.

Future Work

In future research, we will develop personas, including a larger number of participants, thereby allowing to test this combined approach on a larger sample. Moreover, intended use will be coupled with these personas, and usage log data will be used to see whether participants use it as intended. By continuing our research this way, we hope to learn how to attune technological features to our user. We hypothesize that technology personas can inspire developers to put the right persuasive features [29] in the designs and tailor them accordingly to different users. Moreover, in this study, we focused on how users can be better targeted using the PAT method. Specific methods for targeting eHealth are personalization, tailoring, and adapting eHealth. In future research, we aim to carry out a systematic review into how eHealth technologies are personalized. More specifically, we aim to investigate what information from the user is collected to personalize the eHealth technology accordingly. Because we will also map out the effectiveness of these different types of personalization, we can also make a recommendation for the variables that should be considered when developing personas.

Acknowledgments

We acknowledge Dr S Said, A Kleberger, and E Rodijk for clinically supervising the pilot study. Moreover, we want to thank JET Tiebot for contributing to the visual design of the personas.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of the Shapiro-Wilk tests.
[DOCX File , 18 KB - humanfactors_v9i1e24172_app1.docx]

Multimedia Appendix 2
Average silhouette plot for the cluster analysis on the health-related data.

[PDF File (Adobe PDF File), 85 KB - humanfactors_v9i1e24172_app2.pdf]

Multimedia Appendix 3
Meaning of the symbols in the persona descriptions.
[PNG File , 1399 KB - humanfactors_v9i1e24172_app3.png]

Multimedia Appendix 4
Average silhouette plot for the cluster analysis on the health- and person-related data.
[DOCX File , 47 KB - humanfactors_v9i1e24172_app4.docx]

Multimedia Appendix 5
Average silhouette plot for the cluster analysis on the health- and person-related data enriched with usage log data.
[DOCX File , 47 KB - humanfactors_v9i1e24172_app5.docx]

References

1. Peeters JM, Krijgsman JW, Brabers AE, Jong JDD, Friele RD. Use and Uptake of eHealth in General Practice: A Cross-Sectional Survey and Focus Group Study Among Health Care Users and General Practitioners. JMIR Med Inform 2016 Apr 06;4(2):e11 [FREE Full text] [doi: 10.2196/medinform.4515] [Medline: 27052805]
2. Ergonomics of human-system interaction - Part 210: Human-centered design for interactive systems. ISO. 2019 Jul. URL: https://www.iso.org/standard/77520.html [accessed 2022-03-03]
3. van Velsen LS, van Gemert-Pijnen JEWC, Nijland N, Beaujean D, van Steenbergen J. Personas: The linking pin in holistic design for eHealth. 2012 Presented at: The Fourth International Conference on eHealth Telemedicine, and Social Medicine (eTELEMED); January 31; Valencia, Spain p. 128-133 URL: https://research.utwente.nl/en/publications/personas-the-linking-pin-in-holistic-design-for-ehealth
4. Miaskiewicz T, Kozar KA. Personas and user-centered design: How can personas benefit product design processes? Design Studies 2011 Sep;32(5):417-430. [doi: 10.1016/j.destud.2011.03.003]
5. Vosbergen S, Mulder-Wiggers J, Lacroix J, Kemps H, Kraaijenhagen R, Jaspers M, et al. Using personas to tailor educational messages to the preferences of coronary heart disease patients. J Biomed Inform 2015 Feb;53:100-112 [FREE Full text] [doi: 10.1016/j.jbi.2014.09.004] [Medline: 25239261]
6. Haldane V, Koh JJK, Srivastava A, Teo KWQ, Tan YG, Cheng RX, et al. User Preferences and Persona Design for an mHealth Intervention to Support Adherence to Cardiovascular Disease Medication in Singapore: A Multi-Method Study. JMIR Mhealth Uhealth 2019 May 28;7(5):e10465 [FREE Full text] [doi: 10.2196/10465] [Medline: 31140445]
7. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 05;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
8. Rubel P, Fayn J, Nollo G, Assanelli D, Li B, Restier L, et al. Toward personal eHealth in cardiology. Results from the EPI-MEDICS telemedicine project. J Electrocardiol 2005 Oct;38(4 Suppl):100-106. [doi: 10.1016/j.jelectrocard.2005.06.011] [Medline: 16226083]
9. Constitution of the World Health Organization. WHO. 2006. URL: https://www.who.int/governance/eb/who_constitution_en.pdf [accessed 2022-03-03]
10. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
11. van Velsen L, Broekhuis M, Jansen-Kosterink S, Op den Akker H. Tailoring Persuasive Electronic Health Strategies for Older Adults on the Basis of Personal Motivation: Web-Based Survey Study. J Med Internet Res 2019 Sep 06;21(9):11759 [FREE Full text] [doi: 10.2196/11759] [Medline: 31493323]
12. Du Preez V, De La Harpe R. Engaging Aging Individuals in the Design of Technologies and Services to Support Health and Well-Being: Constructivist Grounded Theory Study. JMIR Aging 2019 Mar 20;2(1):e12393 [FREE Full text] [doi: 10.2196/12393] [Medline: 31518258]
13. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: Development of biopsychosocial personas from a study of older adults with heart failure. Int J Med Inform 2017 Dec;108:158-167 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.006] [Medline: 29132622]
14. Ten Klooster I, Noordzij ML, Kelders SM. Exploring How Professionals Within Agile Health Care Informatics Perceive Visualizations of Log File Analyses: Observational Study Followed by a Focus Group Interview. JMIR Hum Factors 2020 Jan 21;7(1):e14424 [FREE Full text] [doi: 10.2196/14424] [Medline: 31961325]
15. Jones JB, Weiner JP, Shah NR, Stewart WF. The wired patient: patterns of electronic patient portal use among patients with cardiac disease or diabetes. J Med Internet Res 2015 Feb 20;17(2):e42 [FREE Full text] [doi: 10.2196/jmir.3157] [Medline: 25707036]

16. Marvin HM. Diseases of the Heart and Blood Vessels: Nomenclature and Criteria for Diagnosis. Arch Intern Med 1964 Jun 01;113(6):906. [doi: 10.1001/archinte.1964.00280120106032]
17. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
18. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
19. Wesselink F, Sieverink F, van Gemert-Pijnen JEWC. The 'Twente TEACH'-pilot: Thales-iMediSense Telemonitoring in Stable Chronic Heart Failure: pilot phase. In: University of Twente. 2016 Presented at: Lecture 'self-monitoring & coaching'; December 19; Enschede URL: https://research.utwente.nl/en/publications/the-twente-teach-pilot-thales-imedisense-telemonitoring-in-stable
20. Fereday J, Muir-Cochrane E. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. International Journal of Qualitative Methods 2006 Mar 01;5(1):80-92. [doi: 10.1177/160940690600500107]
21. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. Int J Endocrinol Metab 2012;10(2):486-489 [FREE Full text] [doi: 10.5812/ijem.3505] [Medline: 23843808]
22. Kaufman P. Partitioning around medoids (program PAM). In: Rousseeuw PJ, editor. Finding Groups in Data: An Introduction to Cluster Analysis. New York: John Wiley & Sons; 1990:68-123.
23. Integrated development environment for R. Team RStudio. 2015. URL: https://www.rstudio.com [accessed 2022-03-03]
24. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. Cluster: cluster analysis basics and extensions. R package version 1 (2). 2012. URL: https://cran.r-project.org/web/packages/cluster/cluster.pdf [accessed 2022-03-03]
25. Wickham H. Elegant graphics for data analysis. In: ggplot2. New York: Springer-Verlag; 2016.
26. Turner P, Turner SE. Triangulation in practice. Virtual Reality 2009 May 12;13(3):171-181. [doi: 10.1007/s10055-009-0117-2]
27. Bruynseels K, Santoni de Sio F, van den Hoven J. Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm. Front Genet 2018;9:31 [FREE Full text] [doi: 10.3389/fgene.2018.00031] [Medline: 29487613]
28. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed J E Health 2012 Mar;18(2):137-144. [doi: 10.1089/tmj.2011.0062] [Medline: 22381060]
29. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: Key Issues, Process Model, and System Features. CAIS 2009;24:485-500. [doi: 10.17705/1CAIS.02428]

Abbreviations

CeHRes: Center for eHealth Research
eHEALS: eHealth Literacy Scale
EPR: electronic patient record
EQ-5D-5L: 5-level 5-dimension Euro quality of life
NYHA: New York Heart Association
PAT: Persona Approach Twente
UCD: user-centered design

Edited by G Eysenbach; submitted 10.09.20; peer-reviewed by M Nitsch, S Santana, D Pförringer; comments to author 14.12.20; revised version received 11.04.21; accepted 15.11.21; published 15.03.22.

Please cite as:
ten Klooster I, Wentzel J, Sieverink F, Linssen G, Wesselink R, van Gemert-Pijnen L
Personas for Better Targeted eHealth Technologies: User-Centered Design Approach
JMIR Hum Factors 2022;9(1):e24172
URL: https://humanfactors.jmir.org/2022/1/e24172
doi:10.2196/24172
PMID:35289759

Original Paper

Development and Usability of a Text Messaging Program for Women With Gestational Diabetes: Mixed Methods Study

Rachel A Blair[1], MD; Christine E Horn[1], BA; Jennifer M Dias[1], BA; Marie E McDonnell[1*], MD; Ellen W Seely[1*], MD

Division of Endocrinology, Diabetes, and Hypertension, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Rachel A Blair, MD
Division of Endocrinology, Diabetes, and Hypertension, Department of Medicine
Brigham and Women's Hospital
Harvard Medical School
221 Longwood Avenue
Boston, MA, 02115
United States
Phone: 1 6177329479
Email: rblair4@bwh.harvard.edu

Abstract

Background: Gestational diabetes mellitus (GDM) affects 5%-10% of pregnancies and can lead to serious fetal and maternal complications. SMS text messaging is an effective way to improve diabetes management outside of pregnancy, but has not been well studied in GDM.

Objective: This study aimed to perform user experience testing and assess usability and acceptability of an SMS text messaging program (Text 4 Success) for women with GDM.

Methods: An automated 2-way texting program was developed. It included (1) reminders to check blood glucose levels, (2) positive feedback to user-reported glucose levels, (3) weekly educational messages, and (4) weekly motivational messages. For the user experience testing, women received simulated messages. For the usability study, women were enrolled in the program and received messages for 2 weeks. All women participated in semistructured interviews. For women in the usability study, data from glucose measuring devices were downloaded to assess adherence to self-monitoring of blood glucose (SMBG), measured as the percentage of recommended SMBG checks performed (a secondary outcome).

Results: Ten women participated in user experience testing. Suggestions for optimization included further customization of message timing and minimization of jargon, which were incorporated. Ten women participated in the usability study. All 10 would recommend the program to other women with GDM. Participants liked the immediate feedback to glucose values. Suggestions included further flexibility of messages related to mealtimes and the ability to aggregate blood glucose data into a table or graph. Overall, adherence to SMBG testing was high at baseline (222/238 recommended checks, 93%). In comparing the week prior to the trial with the 2 weeks during the trial, there was a small but statistically insignificant difference (P=.48) in the percentage of recommended SMBG performed (median 93% [25th-75th IQR 89%-100%] vs median 97% [25th-75th IQR 92%-100%]).

Conclusions: Overall, women with GDM would recommend the Text 4 Success in GDM program and think it is helpful for GDM self-management. The program was usable and acceptable. The program may be better suited to those who have low levels of adherence to SMBG at baseline or to women at time of their diagnosis of GDM. Adaptations to the program will be made based on user suggestions. Further study of SMS text messaging to improve SMBG in GDM is needed.

(JMIR Hum Factors 2022;9(1):e32815) doi:10.2196/32815

KEYWORDS

gestational diabetes mellitus; SMS text messaging; mobile phone; mobile health; pregnancy; blood glucose self-monitoring

Introduction

Gestational diabetes mellitus (GDM) is a common condition, affecting 5%-10% of pregnancies in the United States [1], and has important implications for maternal and child health. Poorly controlled GDM can cause adverse fetal outcomes including preterm delivery, neonatal hypoglycemia, and fetal demise, as well as increased maternal risk for preeclampsia, Cesarean sections, and other complications [2]. The cornerstone of management relies on lifestyle modification and self-monitoring of blood glucose (SMBG), typically 4 times daily. As soon as the diagnosis of GDM is made, women are asked to start intensive monitoring very quickly [2,3]. The blood glucose values obtained not only allow women to understand their blood glucose trends but are also essential for clinicians to tailor therapy, including if pharmacologic treatment is indicated and whether adjustments to dosing are needed. Without the crucial information from SMBG, women and their clinicians cannot work together for optimal glucose control during pregnancy. However, many women have difficulty adhering to this intensive monitoring [4-6], and a study found that women with poor adherence to SMBG are more likely to have poor pregnancy outcomes including preeclampsia [7]. Therefore, a mobile health intervention that could improve SMBG in GDM could be very impactful.

SMS text messaging programs have been shown to improve glycemic control in diabetes outside of pregnancy [8,9]. There is preliminary evidence that SMS text messaging programs are well-received by women with GDM [10,11], though more research is needed. Two-way texting is patient centered and may improve engagement in care but has not been well-studied in GDM. Additionally, the SMS text messaging technology can be applied remotely, which became very relevant during the COVID-19 pandemic [12].

SMS text messaging programs have the advantage of being highly accessible and easily scalable, compared with apps that are only available on smartphones and require downloading and opening for use, and are more expensive to develop [8,13,14]. For these reasons, we developed an automated 2-way SMS text messaging program designed to increase SMBG in women with GDM. This study aimed to first assess user experience (phase 1), followed by usability and acceptability of an SMS text messaging program for women with GDM (phase 2). As a secondary outcome, we aimed to gather data about the program's effect on adherence to SMBG.

Methods

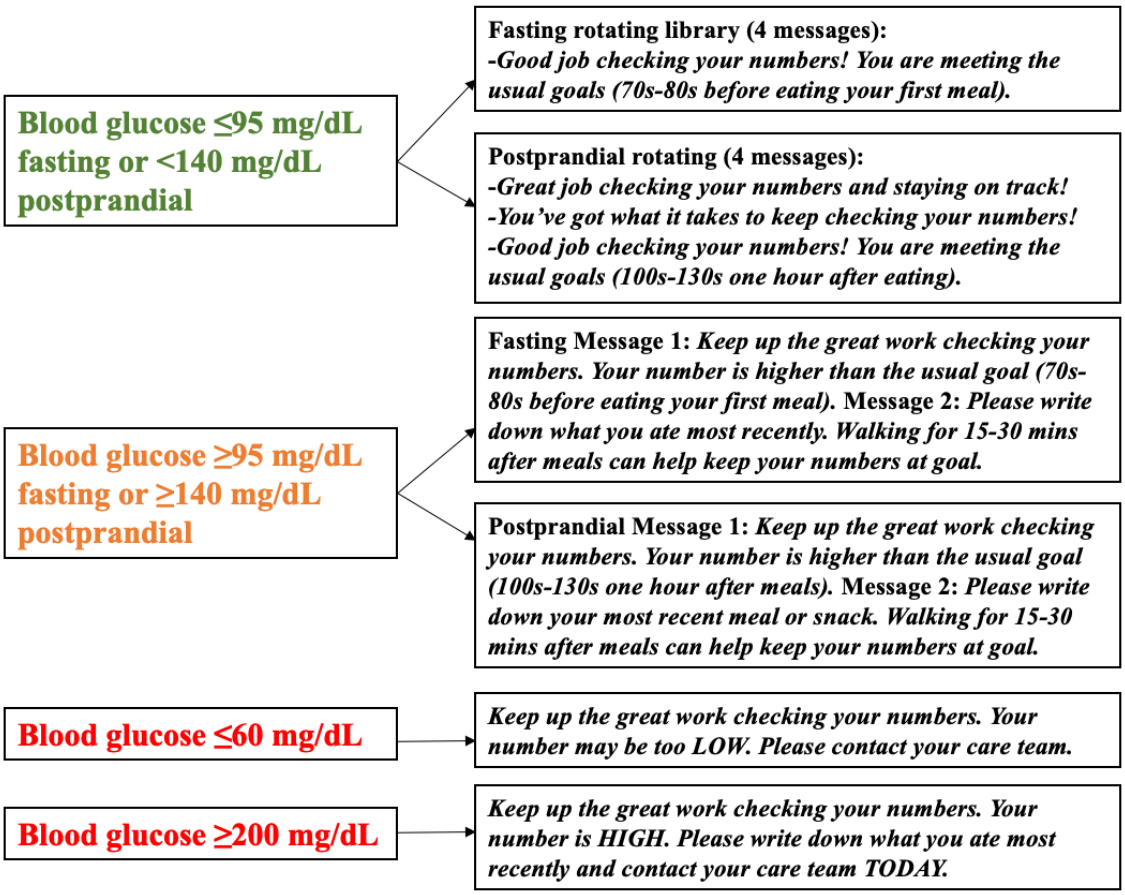
Development of the Text 4 Success in Gestational Diabetes Text Messaging Program

We designed an automated SMS text messaging program for women with GDM called Text 4 Success in Gestational Diabetes (referred to as Text 4 Success). When designing the messages, we applied 2 components of the Health Belief Model: cue-to-action (the stimulus needed to take a health action) and self-efficacy (confidence in one's ability to undertake a health action) [15]. The program included several different types of messages: (1) reminders to check blood glucose values fasting and 1 hour after meals (based on mealtimes reported at enrollment), (2) positive feedback on user-reported blood glucose values (with high or low values automatically prompting users to contact their clinical care team), (3) educational messages, and (4) motivational messages (Table 1). The reminders to check blood glucose values address the cue-to-action component of the Health Belief Model, providing an external stimulus to engage in SMBG [15]. The educational messages were designed to supplement information about GDM, physical activity, and healthy eating provided by clinicians. The positive feedback and motivational messages incorporate elements of self-efficacy from the Health Belief Model [15]. For instance, "You've got what it takes to keep checking your numbers!" specifically relates to a woman's belief in her ability to carry out a behavior. More specific information about each component and sample messages are shown in Table 1 and Figure 1.

Table 1. Text 4 Success in Gestational Diabetes components.

Component	Description	Example
Reminders to check glucose values	Sent 4 times per day, 30 minutes before reported time of breakfast, and then 1 hour after reported time of breakfast, lunch, and dinner. Users were asked to text back their glucose values.	*Hello! This is a Text 4 Success reminder to check your number before eating your first meal. Please reply with your number only.*
Feedback on blood glucose values	In response to submitting glucose values, participants received encouraging feedback. In addition, an automatic algorithm sent messages based on glucose range (Figure 1).	*Keep up the great work checking your numbers and taking charge of your health!*
Educational messages	One per week sent at participants' choice of time per day.	*Eating nutritious foods is a key part of staying healthy. One healthy snack is a plain Greek yogurt. Click for a list of more snacks: http://tinyurl.com/y2az4zn6*
Motivational messages	One per week sent at participants' choice of time per day.	*If you feel off track, know that every day is another chance to get back on track!*

Figure 1. Algorithm for feedback on blood glucose.



The messages were reviewed by a multidisciplinary group of experts, including endocrinologists, maternal–fetal medicine physicians, a behavioral psychologist, and a nutritionist. All information on the texting platform was at an eighth-grade reading level or below and available in English or Spanish based on participant preference. The program was specifically designed as an adjunct to clinical care. Therefore, participants were prompted by the text responses to contact their clinicians for both high and low glucose values that were out of desired range (fasting 61-94 mg/dL, postprandial 61-139 mg/dL). It was also designed in such a way that clinicians would not have the additional responsibility of monitoring the program in real time.

Phase 1: User Experience Testing of the Text 4 Success Text Messaging Program

Participants

Women 18 years and older with GDM were recruited in Boston, Massachusetts, at a tertiary care center and affiliated clinics. Women were recruited by research assistants after clinical appointments as well as by clinicians seeing patients in maternal–fetal medicine obstetric clinics and endocrinology clinics focused on diabetes in pregnancy. A research assistant then obtained verbal informed consent. Participants received a parking voucher worth US $9 for participation.

At this institution, 2-step testing is used for the diagnosis of GDM [2]. GDM was defined by Carpenter–Coustan criteria applied to a 3-hour 100-g oral glucose tolerance test [16], a glucose value of ≥200 mg/dL after a 50-g glucose challenge test at >12 weeks of gestation, or a diagnosis of GDM documented in the electronic health record by a health care provider. Other inclusion criteria included completion of eighth grade, English or Spanish speaking, and ownership of a mobile phone with texting capability. Exclusion criteria included type 1 or type 2 diabetes or a hemoglobin A1c level ≥6.5% in the first trimester.

We performed user experience testing, defined as evaluation of a person's responses as the result of the use of a system [17]. The goal of this process was to gather initial feedback on the design of the program as well as the phrasing and content of the messages in order to incorporate suggested changes.

Study Procedures

Members of our study team (RAB and JMD) met individually with each woman, sent simulated text messages from a study laptop to a study mobile phone held by the participant, and conducted semistructured interviews in English or Spanish to assess their opinions of the program. These meetings were in-person and took approximately 30-45 minutes to conduct. A research assistant (JMD) took notes during each interview. Messages were revised in an iterative manner based on user

feedback. After the changes were made, we launched usability testing.

Phase 2: Usability Testing

Participants

Inclusion criteria were similar to user experience testing. For usability testing, women additionally had to be ≤36 weeks of gestation and needed to have an unlimited SMS text messaging plan. Given that standard of care for GDM includes SMBG, all women were self-monitoring blood glucose typically using a glucometer or, in 1 case, a continuous glucose monitor. Women were recruited by clinicians at routine clinic visits. A research assistant then obtained verbal informed consent. Participants received a parking voucher worth US $22 for completing the study. They did not receive any specific compensation for replying to text messages.

Study Procedures

Women were enrolled in the SMS text messaging program for 2 weeks. The women started receiving text messages 24 hours after enrollment. Participants received a welcome message with opt-out information (users could easily unsubscribe by sending a text message stating “STOP”). They also received a message explaining the account is not monitored by a clinician in real time and to contact their care team with any clinical questions. They then answered a series of brief initial text messages with questions about mealtimes so that their reminders to check blood glucose could be timed accordingly. They were also sent baseline and end of study demographic questionnaires via a message on the secure patient portal used by the institution or by email, depending on patient preference. Questionnaires were designed and administered using REDCap (Research Electronic Data Capture) electronic data capture tools hosted at Harvard Catalyst, the Harvard Clinical and Translational Science Center [18,19]. REDCap is a secure, web-based software platform designed to support data capture for research studies. At the end of 2 weeks, women participated in semistructured interviews so we could obtain their feedback and assess their experience with the program. Interviews took approximately 30-45 minutes to complete and were conducted via Zoom, a cloud-based videoconferencing service offering features including online meetings and secure recording of sessions [20]. Audio-only calls were performed and the recording function was used to record the interviews.

After the study concluded, women were asked to drop off their glucometer or continuous glucose monitor at their clinic sites or mail their glucometer to the study staff using a prepaid shipping method. Data from these devices were then downloaded to determine the number of times that women were checking their blood glucose levels each day 1 week prior to and the 2 weeks during the study. We assessed the electronic health record to determine the number of blood glucose checks per day recommended by the clinician.

Measures and Analyses

Participants reported their age, race/ethnicity, and education level. We accessed the electronic medical records to determine gestational age at diagnosis of GDM as well as gestational age at enrollment into the study.

Likert scale questions were asked to assess usability and acceptability of the program as part of semistructured interviews performed at the end of 2 weeks. Participants were also asked open-ended questions about usability and acceptability of the program to better understand the Likert scale answers, as well as to obtain suggestions for improvement. This method of combining quantitative (Likert scale) and qualitative (open-ended) data has been used in a previous study [21]. Interview recordings were transcribed by an outside transcription service (Landmark Associates, Inc.). Two independent researchers (CEH and RAB) coded responses to determine themes, which were then organized by program component or as general program suggestions. Dedoose (Dedoose, LLC), a qualitative data program, was used to code the interview responses.

As a secondary outcome, we assessed adherence to recommended SMBG for the 1 week prior to enrollment compared with the 2 weeks during enrollment based on the data downloaded from glucose monitoring devices. The percentage of recommended glucose checks was determined by the total number of glucose levels checked divided by the total number of glucose checks recommended by the participant’s clinician. If a participant checked more often than the recommended checks, this percentage was still calculated as 100% of recommended checks.

Statistical Analyses

The adherence rates for each participant were calculated for the week before and 2 weeks during the study. Adherence rates were assessed using the Shapiro–Wilk test for normality. To compare the SMBG adherence the week before and in the 2 weeks during the study, the Wilcoxon rank sum test was used. *P* values of <.05 were considered significant. Median and 25th-75th IQRs are reported.

Ethics Approval

This study was approved by the Partners Healthcare Institutional Review Board, with the protocol numbers 2019P000010 for phase 1 and 2019P002591 for phase 2.

Results

Phase 1: User Experience Testing Results

We performed user experience testing with 10 women with GDM. Women were an average of 33.2 (3.2) years old. Five were White and non-Hispanic Latina, 3 were White and Hispanic/Latina, 1 was Black, and 1 was Asian. Four spoke English as their native language (3 spoke Spanish; 3 spoke other languages [Vietnamese, Haitian Creole, and Greek]). Two user experience sessions were conducted in Spanish using Spanish text messages and interview questions. Seven women had completed a college degree. Three were nulliparous and 3 had GDM in a prior pregnancy. Five were using insulin as treatment for GDM and 5 were on dietary therapy alone.

After the program was explained, all women expressed interest in using the SMS text messaging platform if it were available.

Overall, women preferred messages without medical jargon. For example, our original welcome message was "Welcome to Text 4 Success in Gestational Diabetes (Txt4GDM)! You will receive reminders & info about GDM." An early participant asked what the "M" stood for because GDM is not often referred to as "gestational diabetes mellitus" in conversations with patients. We therefore shortened the name of the program to Text 4 Success. Women suggested customization options, such as the option to pick the time of day to receive educational or motivational messages, which we incorporated. Finally, women wanted the messages to be more specifically related to GDM. For example, an educational message that said "Drinking water, instead of soda or juice, is healthy for you" was edited to include "and can help regulate your numbers" based on participant feedback.

Phase 2: Usability Testing Results

Overview

Ten women underwent usability testing and their characteristics are shown in Table 2. The majority of participants were White and all were college educated. All reported receiving text messages from their doctors' offices or pharmacies. No women replied "STOP" to unsubscribe to the text messages throughout the 2-week study period. On average, women were diagnosed 3 weeks prior to enrollment to the study, during which time they had been performing SMBG. Overall, women responded to 67.9% (380/560) of text messages that they received from the SMS text messaging program.

Table 2. Characteristics of the usability study population (n=10).

Characteristic	Value
Age (years), mean (SD)	36.5 (4.0)
Weeks of gestation at GDM[a] diagnosis, mean (SD)[b]	27.3 (1.1)
Weeks of gestation at study enrollment, mean (SD)	30.4 (5.2)
Race/ethnicity, n (%)	
White	9 (90)
Asian	1 (10)
Hispanic	0 (0)
Black	0 (0)
Multiple	0 (0)
Other	0 (0)
Native language, n (%)	
English	7 (70)
Other[c]	3 (30)
Highest reached education, n (%)	
Some or all of high school	0 (0)
Some college	0 (0)
College graduate	8 (80)
Graduate degree or higher	2 (20)
Nulliparous, n (%)	3 (30)
GDM in prior pregnancy, n (%)	3 (30)
Used insulin during this pregnancy, n (%)	6 (60)

[a]GDM: gestational diabetes mellitus.

[b]Two participants had a clinician diagnosis of GDM and started monitoring blood glucose levels (using fingersticks) early in their pregnancy so we did not have a formal date of diagnosis for them. This value is the average for the other 8 participants.

[c]Other includes Farsi, Gujarati, and Hebrew.

Interview Results

All 10 participants completed the usability interview at the completion of the 2 week study. All participants stated they would recommend the program to other women with gestational diabetes. One participant mentioned that the program may be more helpful for women who do not check their blood glucoses frequently. Four women wanted to use the program for 2 weeks, 1 woman wanted to use it for 1-2 months, 4 wanted to use the program for their entire pregnancy, and 1 wanted flexibility to decide. Eight out of the 10 women thought the amount of text messages was just right (not too many or too few) and 2 thought there were too many messages. Seven of the 10 women said they would prefer an app, either for more flexible timing of

reminders or for a way to aggregate blood glucose levels into a graph or table. One participant was neutral and 2 preferred SMS text messaging.

Responses to the Likert scale questions from all women assessing if the components of the SMS text messaging program were understandable and useful are shown in Table 3, using a Likert scale with 5 as the easiest or most helpful and 1 as not easy/not helpful. For helpfulness of reminders to check blood glucose levels, the mean was 3.3 (moderately helpful), with 5 women rating them as a 4 or 5, 2 women rating them as a 3, and 3 women rating them as a 1 or 2. Participants also reported aspects that they felt could be improved for future iterations. Below we describe the feedback for each component of the SMS text messaging program as well as overall suggestions for the program.

Table 3. User feedback on SMS text messaging components by Likert scale.

Component	Mean (SD)
Four times daily reminders to check glucose	
Understandability[a]	5.0 (0.0)
Helpfulness[b]	3.3 (1.3)
Feedback messages to glucose values	
Understandability	4.8 (0.4)
Helpfulness	3.7 (1.1)
Educational texts	
Understandability	5.0 (0.0)
Helpfulness	3.8 (0.9)
Motivational texts	
Understandability	5.0 (0.0)
Helpfulness	4.0 (1.3)

[a]Participants were asked for each type of message about understandability: on a scale from 1 to 5, where 1 is not easy, 3 is moderately easy, and 5 is very easy, how easy was it for you to understand these messages?

[b]Participants were asked for each type of message about helpfulness: On a scale from 1 to 5, where 1 is not helpful, 3 is moderately helpful, and 5 is very helpful, how helpful were these messages?

Reminders to Check Glucose Levels

Participants felt that reminders to check glucose levels, which included a request to text back in a glucose value, were easy to understand and helpful (Table 3). Participants liked that the reminders were brief and clear. One woman spoke specifically about how challenging it can be to suddenly have to engage in intensive monitoring and how the reminders were helpful in making that transition.

> *I think it can be overwhelming to start having to check from going to not doing this at all and then having to do it four plus times per day. Particularly for people who are really busy, it's just so easy to forget. Even since I stopped getting the text messages, today I remembered it'd been two and a half hours since I last took my blood sugar. I think just getting those reminders made it a lot easier to remember to check.*

Participants' mealtimes varied day to day so they would have preferred more flexibility in being able to text in numbers at other times of the day rather than the set mealtimes entered at the beginning of the program.

> *It would almost work better if I could text in and say, "Just finished eating,"...and then it would give you a reminder that an hour later to test.*

Two women suggested that if they did not reply to a reminder with their glucose value, a second reminder be sent after a given period.

Feedback to Glucose Values

Participants liked that the program had 2-way SMS text messaging and that feedback messages were sent back to them based on glucose values sent to the program. In addition, 2 women used the SMS text messaging chain as a way to track their glucoses, either transcribing them to tracking sheets or as a primary way to report them to their providers.

> *It was very nice that as soon as you input, it immediately tells you whether you're in range or not. That's super nice. If you don't know what you actually have to be it will tell you you're doing good or no, you should be in this range.*

Four women found the text of the replies to glucose values repetitive and suggested more variety in the feedback responses. There were 4 rotating replies available when glucoses were in normal range, and only 1 reply if low (<60 mg/dL) or high (≥200 mg/dL).

> *I think having a wider range of responses that you get would be great. Some of the responses tended to be the same each day.*

Educational Messages

Women appreciated practical educational tips, especially the text message that had a link to healthy snacks. Overall, 7 women wanted more than once weekly messages with educational content. Three suggested that the program could include more texts with links to additional information, such as exercise videos, healthy dessert recipes, or advice specifically around high fasting blood glucose.

> *Having only two messages in that week was like, okay, what else is there to know? I feel like there's gotta be more, right? Especially maybe even frontloading more information early in the program would be nice.*

Motivational Messages

Seven women found these messages quite encouraging and liked them.

> *When you get a discouraging number, it's easy to get stuck in your head, so to just get a positive reinforcement like, "It's just one number. You can get back on track," it was encouraging.*

One woman did not find the motivational messages particularly helpful and 2 women were neutral. One woman suggested that the frequency of motivational messages could be adjusted based on user preferences, which could allow for women to select these types of messages more or less often depending on how helpful they found them.

General Program Suggestions

Overall, suggestions fell into 2 major categories: increased functionality and increased customizability. In terms of increased functionality, 7 participants suggested that there be some sort of graph or table that could aggregate their responses to the reminders.

> *It would be really nice if it was something that would generate a weekly report or something about your numbers since it's taking that information that then you could share with the doctor or maybe could be automatically shared with your physician. I think something like that would be really, really helpful in, again, accountability and having a way to know when you're overall on track or off track.*

This type of graph or table functionality could either be achieved using an accompanying website associated with the texting platform or via an app. Specifically related to educational content, 1 woman suggested an app as that could have a more comprehensive library of information.

Second, women reported wanting increased customizability of the program. For example, 2 women specifically mentioned that they would have preferred to pick different mealtimes on weekdays compared with weekends. Two suggested that the educational messages could be tailored to areas of GDM management they were most struggling with, such as receiving educational messages specifically about managing fasting glucose levels.

Self-Monitoring of Blood Glucose Adherence Results

All participants had their glucose monitoring data downloaded from their glucose measuring devices (9 glucometers and 1 continuous glucose monitor).

We assessed the adherence to SMBG by calculating the percentage of recommended checks. The data for 1 week prior to the study were available for 9 out of 10 women because 1 woman was diagnosed and started SMBG at time of enrollment. Women were checking a median of 93% (25th-75th IQR 89%-100%) of recommended fingerstick blood glucose levels in the week prior to enrollment. During the 2 weeks in the trial, women checked a median of 97% (25th-75th IQR 92%-100%) of recommended glucose levels. The percentage of recommended fingerstick blood glucose levels was not significantly different when comparing the 1 week prior with the 2 weeks during the study (P=.48).

Discussion

Principal Findings

Overall, the Text 4 Success in Gestational Diabetes program was found to be usable and acceptable to women with GDM, with all women in the usability study saying they would recommend the program to other women with GDM. We used 2 components of the Health Belief Model to design the program: cue-to-action and self-efficacy [15]. The cue-to-action reminders were helpful according to the women in the study. The motivational messages targeting self-efficacy were well-received. The first step of user experience testing allowed us to refine messages and program structure prior to the usability study in an iterative fashion. In the 2-week usability study, 8 of 10 women thought the number of text messages (up to 4 reminders and 4 feedback messages per day) was just right. This number of text messages is higher than is usually seen for SMS text messaging programs during pregnancy (3-7 messages per week) [10,11,22]. There was a range in how long women wanted to use the program, with 4 reporting they would want to use it for their entire pregnancy.

Women reported that the program helped them make the rapid transition to SMBG 4 times daily, which is typically required in GDM. With high acceptability of the program, women were interested in further adaptations to the program, such as more flexible mealtimes, the ability to aggregate blood glucose values into a chart or graph, and more educational content. Although many stated they would prefer an app, some of the reasons they wanted an app could be addressed using a modified text message program alone or with an accompanying website. For example, for more flexible mealtimes, we could follow the suggestion of having users text in a keyword or keywords such as "ate meal" and then users would get a text reminder 1 hour later to check their blood glucoses. Additionally, an SMS text messaging program could have an accompanying website with a graph or table of glucose values and could include additional information about GDM as well. In addition to patients being able to see their blood glucose levels, clinicians could also view the data via a secure portal.

Alternatively, an app could be developed with push notifications for reminders to check blood glucose levels, more educational content, and summaries of glucose data. In fact, several apps for GDM have been developed without the specific focus of SMBG adherence. These apps have a range of features including graphs of glucose data, though few include specific reminders to check blood glucose levels [23,24]. When considering a future iteration of this program, it is notable that the cost of developing and maintaining an app is much more resource intensive than developing and maintaining an SMS text messaging program [14], and is thus more expensive [25,26]. Additionally, in resource-poor settings, fewer patients have access to smartphones with the ability to use apps [27]. A median of 76% of people in advanced economies own a smartphone and a median of 45% of people in emerging economies own a smartphone, whereas 94% of people in advanced economies and 83% of people in emerging economies own a mobile phone [27].

The study was not powered to detect differences in SMBG adherence comparing baseline with study duration and the intervention did not significantly increase the adherence rate for SMBG. Many women in the usability study had a very high adherence rate to SMBG (a median of 93%) prior to the study, which is higher than the average of approximately 70% adherence rate that has been described [28]. This high baseline adherence rate suggests that in the future, the program would be better suited to patients who have a lower baseline adherence to SMBG. The participants in the usability study all had a high level of education (college degree or higher) which may have played a role in their high baseline adherence rate [7].

Limitations

There were several limitations to our study. The study was conducted at a single academic center. Participants were recruited by clinicians, so we do not have characteristics of those who declined participation. Participants in the usability testing all had a college degree or higher which could make findings less generalizable to patients of a lower socioeconomic status. The messages were well received in the user experience testing group as well, which had 3 participants without a college degree. The usability study only included 10 women, though other usability studies can have similar sample sizes [29,30]. The usability study did not include Spanish-speaking women. The usability testing only lasted 2 weeks so we do not have opinions on how the intervention would be received for a longer time frame. Participants only received 2 educational and 2 motivational text messages in the 2-week study, so feedback on those components is limited. The usability testing itself was not iterative, though it built upon feedback from the user experience testing. Finally, there was approximately a 3-week period between diagnosis of GDM and enrollment into the study, which may have led to the high baseline rates of glucose monitoring.

Comparison With Prior Work

There have been a few recent studies assessing the impact of automated messaging on adherence to SMBG [11,31]. Johnson et al [11] conducted a 4-week intervention with 1-way texting in 19 women with GDM in the United States. One text message per day was sent, either a reminder to check blood glucose or an educational message. Nearly 67% felt the messages helped them remember to check glucose levels. In contrast to Johnson et al [11], our program gave immediate feedback to blood glucose levels via an algorithm, which women in our study reported that they found helpful.

In contrast to using an SMS text messaging program, several studies have used an app to increase SMBG. Peleg et al [31] designed an app that sent messages to user smartphones to encourage monitoring of a variety of different parameters including blood glucose in 19 participants with GDM in Spain. Glucometers were connected to user smartphones via Bluetooth and 4 reminders to check blood glucose levels were sent daily based on entered mealtimes. The system sent a message in response to elevated blood glucose readings to the patient and to the care provider. There was an improvement in mean adherence to SMBG in the intervention group (101%, SD 10%) compared with mean adherence in a historical control of 247 patients (87%, SD 28%; P=.03). Adherence was calculated such that it could be >100%. Neither Johnson et al [11] nor Peleg et al [31] described basing their program on an underlying health behavior change theory and neither mentioned soliciting patient input in the message development process.

There have been 3 randomized controlled trials of interventions conducted in Israel, China, and the United Kingdom that involved frequent communication with clinicians facilitated by an app and evaluated the effect on compliance to glucose monitoring [28,32,33]. Data on glucose levels were transmitted via Bluetooth from glucometers (or could be manually entered into the app in 1 study [32]). These 3 studies were similar to one another in that all required intensive communication from clinicians to participants (either daily or 3 times per week). None of the studies mentioned how these responsibilities were balanced with other clinical care. All showed improvement not only in adherence to SMBG but also in glycemic control.

There was one recent study examining a 1-way SMS text messaging program in GDM that was not specifically related to SMBG adherence [10]. The program sent 3 supportive or educational messages per week. Participants felt that the messages helped their motivation for diabetes self-care. Similar to our findings, participants wanted more educational and supportive messages and also desired more recipes.

Conclusions

Our program is a novel 2-way texting program designed for women with GDM consisting of automated reminders and feedback to patients about their blood glucose values without requiring clinical staff to manage messages in real time. It allows feedback to be given by an algorithm rather than using clinician time, which has been brought up as a criticism of 2-way texting scalability [8]. Two components of the Health Belief Model (cue-to-action and self-efficacy), along with patient input, were used to design and refine the program. The program was easily understood and well received. The program may be better suited to women with a low baseline adherence rate to SMBG or to women at the time of their diagnosis of GDM. Women provided suggestions to improve the program, including having more customizability and functionality, which could be achieved with

an accompanying website or by conversion to an app. These suggestions will be incorporated into the next iteration of the intervention. Further study, including a randomized controlled trial, is needed to assess this SMS text messaging program on adherence to SMBG.

Acknowledgments

RAB was supported by NIH/NIDDK (National Institutes of Health/National Institute of Diabetes and Digestive and Kidney Diseases) F32 DK124964-01. RAB also received an Endocrine Fellows Foundation grant and a Partners Center of Expertise Health Policy grant to support the project. The authors thank Drs Nadine Palermo and Susan Haden for early feedback on the program and assistance with recruitment. The authors also thank Mahyar Heydarpour for statistical advice and assistance with analysis. This work was conducted with support from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Advancing Translational Sciences, National Institute of Health Award UL1 TR002541) and financial contributions from Harvard University and its affiliated academic health care centers. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, or the National Institutes of Health.

Authors' Contributions

RAB designed the SMS text messaging program and wrote the first draft of the manuscript. RAB and JMD performed user experience testing. RAB and CEH performed usability testing and coded the semistructured interviews. MEM provided input on development of the program, assisted with recruitment, and oversaw download of glucose measuring devices. EWS oversaw the study design, its completion, and the writing of the manuscript. All authors reviewed and approved the manuscript for publication.

Conflicts of Interest

None declared.

References

1. DeSisto CL, Kim SY, Sharma AJ. Prevalence estimates of gestational diabetes mellitus in the United States, Pregnancy Risk Assessment Monitoring System (PRAMS), 2007-2010. Prev Chronic Dis 2014 Jun 19;11:E104 [FREE Full text] [doi: 10.5888/pcd11.130415] [Medline: 24945238]
2. ACOG Practice Bulletin No. 190: Gestational Diabetes Mellitus. Obstet Gynecol 2018 Feb;131(2):e49-e64. [doi: 10.1097/AOG.0000000000002501] [Medline: 29370047]
3. Garrison A. Screening, diagnosis, and management of gestational diabetes mellitus. Am Fam Physician 2015 Apr 01;91(7):460-467 [FREE Full text] [Medline: 25884746]
4. Saha S. Compliance and barriers to self-monitoring of blood glucose in patients with gestational diabetes mellitus: A systematic review. Int J Health Sci (Qassim) 2019;13(3):44-52 [FREE Full text] [Medline: 31123440]
5. Guo J, Long Q, Li J, Wang X, Li Y, Jiang S, et al. Barriers and facilitators of self-monitoring of blood glucose engagement among women with gestational diabetes mellitus in China: A mixed-methods study. Midwifery 2020 Nov;90:102797 [FREE Full text] [doi: 10.1016/j.midw.2020.102797] [Medline: 32755742]
6. Martis R, Brown J, McAra-Couper J, Crowther CA. Enablers and barriers for women with gestational diabetes mellitus to achieve optimal glycaemic control - a qualitative study using the theoretical domains framework. BMC Pregnancy Childbirth 2018 Apr 11;18(1):91 [FREE Full text] [doi: 10.1186/s12884-018-1710-8] [Medline: 29642898]
7. Cosson E, Baz B, Gary F, Pharisien I, Nguyen MT, Sandre-Banon D, et al. Poor Reliability and Poor Adherence to Self-Monitoring of Blood Glucose Are Common in Women With Gestational Diabetes Mellitus and May Be Associated With Poor Pregnancy Outcomes. Diabetes Care 2017 Sep;40(9):1181-1186. [doi: 10.2337/dc17-0369] [Medline: 28724718]
8. Haider R, Sudini L, Chow CK, Cheung NW. Mobile phone text messaging in improving glycaemic control for patients with type 2 diabetes mellitus: A systematic review and meta-analysis. Diabetes Res Clin Pract 2019 Apr;150:27-37. [doi: 10.1016/j.diabres.2019.02.022] [Medline: 30822496]
9. Saffari M, Ghanizadeh G, Koenig HG. Health education via mobile text messaging for glycemic control in adults with type 2 diabetes: a systematic review and meta-analysis. Prim Care Diabetes 2014 Dec;8(4):275-285. [doi: 10.1016/j.pcd.2014.03.004] [Medline: 24793589]
10. Yee L, Taylor S, Young M, Williams M, Niznik C, Simon M. Evaluation of a Text Messaging Intervention to Support Self-Management of Diabetes During Pregnancy Among Low-Income, Minority Women: Qualitative Study. JMIR Diabetes 2020 Aug 10;5(3):e17794 [FREE Full text] [doi: 10.2196/17794] [Medline: 32773367]
11. Johnson QB, Berry DC. Impacting diabetes self-management in women with gestational diabetes mellitus using short messaging reminders. J Am Assoc Nurse Pract 2018 Jun;30(6):320-326. [doi: 10.1097/JXX.0000000000000059] [Medline: 29878964]
12. Watson AR, Wah R, Thamman R. The Value of Remote Monitoring for the COVID-19 Pandemic. Telemed J E Health 2020 Sep;26(9):1110-1112. [doi: 10.1089/tmj.2020.0134] [Medline: 32384251]

13. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
14. Turner-McGrievy GM, Hales SB, Schoffman DE, Valafar H, Brazendale K, Weaver RG, et al. Choosing between responsive-design websites versus mobile apps for your mobile behavioral intervention: presenting four case studies. Transl Behav Med 2017 Jun;7(2):224-232 [FREE Full text] [doi: 10.1007/s13142-016-0448-y] [Medline: 27812798]
15. Champion VL, Skinner CS. The health belief model. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass; 2008:45-65.
16. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. Am J Obstet Gynecol 1982 Dec 01;144(7):768-773. [doi: 10.1016/0002-9378(82)90349-0] [Medline: 7148898]
17. Vermeeren A, Law E, Roto V, Obrist M, Hoohout J, Vaananen-Vaino-Mattila K. User experience evaluation methods: current state and development needs. 2010 Presented at: NordiCHI '10: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries; October 16-20, 2010; Reykjavik, Iceland. [doi: 10.1145/1868914.1868973]
18. 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]
19. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
20. Zoom Video Communications. Zoom, Version 5.7.6 (1320). 2020. URL: https://zoom.us/ [accessed 2022-02-11]
21. Nelson LA, Mayberry LS, Wallston K, Kripalani S, Bergner EM, Osborn CY. Development and Usability of REACH: A Tailored Theory-Based Text Messaging Intervention for Disadvantaged Adults With Type 2 Diabetes. JMIR Hum Factors 2016 Sep 08;3(2):e23 [FREE Full text] [doi: 10.2196/humanfactors.6029] [Medline: 27609738]
22. Whittaker R, Matoff-Stepp S, Meehan J, Kendrick J, Jordan E, Stange P, et al. Text4baby: development and implementation of a national text messaging health information service. Am J Public Health 2012 Dec;102(12):2207-2213. [doi: 10.2105/AJPH.2012.300736] [Medline: 23078509]
23. Chen Q, Carbone ET. Functionality, Implementation, Impact, and the Role of Health Literacy in Mobile Phone Apps for Gestational Diabetes: Scoping Review. JMIR Diabetes 2017 Oct 04;2(2):e25 [FREE Full text] [doi: 10.2196/diabetes.8045] [Medline: 30291088]
24. Zahmatkeshan M, Zakerabasali S, Farjam M, Gholampour Y, Seraji M, Yazdani A. The use of mobile health interventions for gestational diabetes mellitus: a descriptive literature review. J Med Life 2021;14(2):131-141 [FREE Full text] [doi: 10.25122/jml-2020-0163] [Medline: 34104235]
25. Kuligowski K. The Best Text Message Marketing Services of 2021. Business News Daily. 2022 Jan 25. URL: https://www.businessnewsdaily.com/15044-best-text-message-marketing-solutions.html [accessed 2021-06-22]
26. Lastovestska A. App Development Cost: Understand Your Budget To Build Powerful Apps. MLSDev. 2021 Nov 12. URL: https://mlsdev.com/blog/app-development-cost [accessed 2021-08-03]
27. Silver L. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Pew Research Center. 2019 Feb 05. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2021-06-22]
28. Miremberg H, Ben-Ari T, Betzer T, Raphaeli H, Gasnier R, Barda G, et al. The impact of a daily smartphone-based feedback system among women with gestational diabetes on compliance, glycemic control, satisfaction, and pregnancy outcome: a randomized controlled trial. Am J Obstet Gynecol 2018 Apr;218(4):453.e1-453.e7. [doi: 10.1016/j.ajog.2018.01.044] [Medline: 29425836]
29. Hsieh K, Fanning J, Frechette M, Sosnoff J. Usability of a Fall Risk mHealth App for People With Multiple Sclerosis: Mixed Methods Study. JMIR Hum Factors 2021 Mar 22;8(1):e25604 [FREE Full text] [doi: 10.2196/25604] [Medline: 33749609]
30. Bhattacharjya S, Cavuoto LA, Reilly B, Xu W, Subryan H, Langan J. Usability, Usefulness, and Acceptance of a Novel, Portable Rehabilitation System (mRehab) Using Smartphone and 3D Printing Technology: Mixed Methods Study. JMIR Hum Factors 2021 Mar 22;8(1):e21312 [FREE Full text] [doi: 10.2196/21312] [Medline: 33749608]
31. Peleg M, Shahar Y, Quaglini S, Broens T, Budasu R, Fung N, et al. Assessment of a personalized and distributed patient guidance system. Int J Med Inform 2017 May;101:108-130. [doi: 10.1016/j.ijmedinf.2017.02.010] [Medline: 28347441]
32. Guo H, Zhang Y, Li P, Zhou P, Chen L, Li S. Evaluating the effects of mobile health intervention on weight management, glycemic control and pregnancy outcomes in patients with gestational diabetes mellitus. J Endocrinol Invest 2019 Jun;42(6):709-714. [doi: 10.1007/s40618-018-0975-0] [Medline: 30406378]
33. Mackillop L, Hirst JE, Bartlett KJ, Birks JS, Clifton L, Farmer AJ, et al. Comparing the Efficacy of a Mobile Phone-Based Blood Glucose Management System With Standard Clinic Care in Women With Gestational Diabetes: Randomized Controlled Trial. JMIR Mhealth Uhealth 2018 Mar 20;6(3):e71 [FREE Full text] [doi: 10.2196/mhealth.9512] [Medline: 29559428]

Abbreviations

GDM: gestational diabetes mellitus
SMBG: self-monitoring of blood glucose

Edited by A Kushniruk; submitted 16.08.21; peer-reviewed by L Nelson, M Zahmatkeshan, N Larburu; comments to author 02.10.21; revised version received 29.10.21; accepted 09.12.21; published 22.02.22.

Please cite as:
Blair RA, Horn CE, Dias JM, McDonnell ME, Seely EW
Development and Usability of a Text Messaging Program for Women With Gestational Diabetes: Mixed Methods Study
JMIR Hum Factors 2022;9(1):e32815
URL: https://humanfactors.jmir.org/2022/1/e32815
doi:10.2196/32815
PMID:35191851

Original Paper

User-Centered Design to Enhance mHealth Systems for Individuals With Dexterity Impairments: Accessibility and Usability Study

Kuntal Chowdhary[1], MD; Daihua Xie Yu[2], PhD; Gede Pramana[2], PhD; Matthew Mesoros[3], BS; Andrea Fairman[4], PhD, MOT, OTR/L, CPRP; Brad Edward Dicianno[1,3], MD; Bambang Parmanto[2], PhD

[1]Department of Physical Medicine and Rehabilitation, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

[2]Department of Health Information Management, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA, United States

[3]Human Engineering Research Laboratories, Department of Veterans Affairs, Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, PA, United States

[4]Occupational Therapy Doctorate Program, Johnson & Wales University, Providence, RI, United States

Corresponding Author:
Bambang Parmanto, PhD
Department of Health Information Management
School of Health and Rehabilitation Sciences
University of Pittsburgh
6052 Forbes Tower
Pittsburgh, PA, 15260
United States
Phone: 1 412 383 6649
Email: parmanto@pitt.edu

Abstract

Background: Mobile health systems have been shown to be useful in supporting self-management by promoting adherence to schedules and longitudinal health interventions, especially in people with disabilities. The Interactive Mobile Health and Rehabilitation (iMHere) system was developed to empower people with disabilities and those with chronic conditions with supports needed for self-management and independent living. Since the first iteration of the iMHere 1.0 app, several studies have evaluated the accessibility and usability of the system. Potential opportunities to improve and simplify the user interface were identified, and the iMHere modules were redesigned accordingly.

Objective: In this study, we aim to evaluate the usability of the redesigned modules within the iMHere 1.0 app.

Methods: We evaluated the original and redesigned iMHere modules—MyMeds and SkinCare. The Purdue Pegboard Test was administered to assess the participants' dexterity levels. Participants were then asked to perform a set of tasks using both the original and redesigned MyMeds and SkinCare modules to assess their efficiency and effectiveness. Usability was measured using the Telehealth Usability Questionnaire to evaluate 10 new accessibility features that were added to the redesigned app. Participants were also asked which version they preferred.

Results: In total, 24 participants with disabilities and varying degrees of dexterity impairments completed the entire study protocol. Participants displayed improved efficiency and effectiveness when using the redesigned modules compared with the original modules. The participants also reported improved usability and preferred the redesigned modules.

Conclusions: This study demonstrated that the iMHere system became more efficient, effective, and usable for individuals with dexterity impairments after redesigning it according to user-centered principles.

(JMIR Hum Factors 2022;9(1):e23794) doi:10.2196/23794

KEYWORDS

cellular phone; mobile apps; telemedicine; adaptive mHealth; rehabilitation; self-care; spinal cord injury; spina bifida; chronic disease; persons with disability; accessibility; dexterity impairments; mobile phone

Introduction

Background

The advent of smartphones has transcended the mobile phone's original purpose—the ability to make phone calls anywhere. Notably, smartphones have radically altered the way people communicate with friends and family, coordinate daily activities, and organize their lives. At the most fundamental level, smartphone users expect their devices to provide an immediate and reliable means of managing their everyday lives [1,2].

One of the most significant emerging trends in the health-related use of smartphones is the proliferation of mobile health (mHealth) apps. These apps can be implemented in a variety of settings, with many focusing on monitoring, managing, and supporting health-related behavior changes [3]. One of the most common type of health-related app focuses on the management of chronic conditions, such as obesity, chronic pain, and type 2 diabetes mellitus, through patient empowerment [4-6].

People with disabilities, however, are one of the largest populations facing health issues that limit their function and participation. The World Health Organization estimates that over 1 billion people, about 15% of the world's population, live with some form of disability [7]. As the population continues to age, the rate of disability continues to rise, in part owing to chronic conditions and the effects of aging itself. Many people with disabilities also have limited access to health care.

Given the high degree of portability and adaptability, mHealth can facilitate self-management and community integration by providing support when the user is between medical visits or in any location, including outside the home. These features may be particularly useful in supporting people with disabilities, who often have limited access to health care and community-based resources to support independent living. The support provided by mHealth may allow users to address secondary complications, which are not always addressed adequately in the outpatient setting, thereby reducing the cost of care [8-10]. Strong evidence supports the importance of using tools to promote self-management skills to improve the health outcomes and independence of people with disabilities [11,12].

Despite the need for mHealth tools to support self-management, a Pew Research Center survey in 2016 found that 65% of people with disabilities have low confidence in their ability to use the internet and other communication devices to keep up with information [13]. This is further compounded by a general lack of awareness of the accessibility features of apps and the skills to use mobile devices optimally [14]. In addition, many mainstream mHealth apps are not designed to address usability or accessibility [15].

The Interactive Mobile Health and Rehabilitation (iMHere) system is an mHealth system that was developed to empower people with disabilities and those with chronic conditions with the skills needed for self-management and independent living [16]. The iMHere 1.0 system originally consisted of a smartphone app for people with disabilities and a web-based portal for clinicians, bridged by a 2-way communication protocol (Figure 1).

Figure 1. Interactive Mobile Health and Rehabilitation platform—MyMeds and SkinCare modules as seen by user.

The iMHere 1.0 app comprised a suite of 5 modules to support medication management (MyMeds), skin integrity (SkinCare), bowel management, bladder self-catheterization, and mental health. People with disabilities could use this suite of modules to report compliance with treatment regimens, ask questions, and receive personalized treatment plans, educational materials, and messages from the clinician. On the clinician side, a web-based monitoring portal allowed clinicians to engage with patients and track their adherence to a specific and individualized treatment plan. By accessing the iMHere portal,

clinicians were able to monitor patients' adherence to self-management activities, view reported problems and issues, and send personalized treatment plans to patients [16].

Given the vast health care implications of using mHealth solutions in people with disabilities, usability testing of mHealth apps is needed. Usability testing has been widely used in the people with disabilities population to test mobile self-management programs. Payne et al [17] demonstrated that usability testing of a web-based e-counseling platform to promote behavioral self-management in patients with chronic heart failure had favorable outcomes in improving the navigation of the website. Williams et al [18] also assessed the usability of a pediatric cardiovascular disease risk factor tool, which yielded revisions through tester feedback to make the mobile app more user-friendly. Thirumalai et al [19] evaluated the development process of a telehealth app used by people with multiple sclerosis through a usability study, which incorporated revisions into the final app. These previous works highlight the importance of usability testing, as it can help identify issues specific to the people with disabilities population that may not have been addressed by program developers in the first iteration of the mHealth solution.

We conducted extensive user acceptance and usability testing of the iMHere system. In the past, 3 studies on the accessibility of iMHere 1.0 have been conducted. In the first study by Yu et al [20], the iMHere 1.0 system was tested for usability. In this study, the modules were tested for self-management workflow, user interface and navigation, and patient-clinician communication. All participants in the study were interested in daily use of the phone app, with the MyMeds and SkinCare modules used frequently by all users, as demonstrated by the consistent use of the phone app during the 6-month intervention period. The clinical portal allowed clinicians to continually monitor patients' conditions and take appropriate steps to prevent secondary complications [20].

In a subsequent study by Yu et al [21], the iMHere 1.0 app was tested for accessibility in 6 participants with spina bifida (SB). The study specifically explored participant experiences with the user interface and the navigation of the module. All 6 participants viewed the modules positively with regard to their support for self-management activities, as indicated by the Telehealth Usability Questionnaire (TUQ) scores (6.52/7 points, 93%). This was further strengthened by the efficiency of performance, as it was noted that shorter times to complete tasks and reduced error rates were seen over repeated trials. In this study, a few avenues for improvement to accessibility were identified, including the need for changes to accommodate users with dexterity impairments.

In a subsequent study, Yu et al [22] explored the accessibility needs and preferences of iMHere users with various disabilities that lead to dexterity impairments. Participants completed 5 tasks, and the *difficulty-on-performance* (DP) was calculated. As expected, a higher degree of dexterity impairment demonstrated more problems in task completion. A few potential issues and barriers were identified, including changes needed to the user interface to create a consistent design, instructive guidance, and simpler cognitive processes in the use of the app.

Objectives

The modules within iMHere 1.0 were redesigned based on these prior studies. The aim of this study is to evaluate the usability of redesigned modules within iMHere 1.0. Hypothesis 1 was that usability (as defined by efficiency and effectiveness) would be higher when completing tasks in the redesigned modules compared with the original modules. Hypothesis 2 was that usability (as measured by the TUQ, which evaluates learnability and satisfaction) would be higher in the redesigned modules, compared with the original modules [23].

Methods

Overview

Modules

This study was designed to evaluate the usability of two modules of the iMHere system: the original and redesigned versions of MyMeds and SkinCare. These modules were specifically selected because of the high rates of medication use and pressure ulcers in the people with disabilities population. Medication mismanagement and inadequate care of pressure injuries are the causes of high rates of hospitalization in the people with disabilities population and significantly increase morbidity [24-26]. These modules were also the most complex iMHere modules in terms of functionality.

MyMeds Module

The MyMeds module helps users manage their medications by providing reminders and monitoring medication adherence. Persons with conditions, such as SB and spinal cord injuries (SCIs), for example, are frequently prescribed several medications for managing neurogenic bowel, neurogenic bladder, spasticity, pain, and depression. Taking multiple medications multiple times per day, while at the same time having to consistently follow other complex self-management routines can be particularly challenging. The MyMeds module helps patients adhere to their medication regimens by providing reminders or cues, keeping track of all their medications and medication schedules (including those medications currently prescribed or prescribed in the past), and reporting if and when the medications have been taken.

SkinCare Module

The SkinCare module reminds users to perform routine inspections of their skin, enables users to take pictures and track any wound or skin conditions that have developed, and at the same time provides the ability to communicate with clinicians on how to care for these problems. For people with SB or SCI; for example, constant vigilance is needed to prevent pressure injuries, particularly in the lower limbs and buttocks, where sensation may be impaired. Pressure injuries are not only detrimental to the patient owing to increased mortality and increased intensive care unit and hospital length of stay but also present a significant health care burden given increased health care costs and health care use following discharge [27-29].

Study Design

This study was approved by the Institutional Review Board of the University of Pittsburgh (PRO12090453). All participants provided informed consent for participation. The participants were recruited from local outpatient rehabilitation medicine clinics. A sample size calculation was performed using the Wilcoxon signed-rank test (2-sided). A sample size of 14 achieved 91% power to detect a mean of paired differences of 1.0, with an estimated SD of paired differences of 1.0, with a significance level (α) of .05.

The inclusion criteria were as follows: users must be between the ages of 18 and 64 years, have fine motor dexterity impairments, have potential for skin breakdown (defined by diagnosis or lack of sensation), and use at least one (prescription or nonprescription) medication. The exclusion criteria were as follows: users with vision, hearing, or speech limitations that entirely precluded the use of a smartphone. Individuals were not excluded if they had used iMHere in a prior study, but a 4-month washout period was used to mitigate learning effects.

Usability was defined according to the usability attributes by Nielsen [30]. The Nielsen model of usability was selected as a framework for this study, as it is multifaceted in its approach to the many dimensions of usability. We examined the usability constructs of efficiency and effectiveness (including errors; hypothesis 1) by assessing task time and errors made. We also used a validated usability survey (TUQ) to measure the usability constructs of learnability and satisfaction (hypothesis 2). We also evaluated user preferences. This design has been used in prior research [31-34]. Participants were first randomly assigned to use either the original or redesigned modules. Participants were then crossed over and provided with alternate modules, such that each participant served as his or her own matched control. As such, we elected not to test memorability in this study, as testing of memorability would confound our washout period between testing of the original and redesigned modules. Data were collected either in the laboratory or at the site of the participant's choosing (ie, home or office).

Demographics, Training, and Dexterity

A background questionnaire was administered to collect the participants' demographic data, previous experience with mobile phones, and knowledge of mHealth technologies.

A face-to-face orientation and training session (approximately 15 minutes) was conducted to introduce the MyMeds and SkinCare modules. Participants were trained to perform the tasks for each of the modules using a trial medication bottle and a mock skin problem image. Participants were scheduled to complete the second set of modules after a 3-week period. This crossover period served as the washout period to minimize the learning effects.

To assess the participants' dexterity levels, the Purdue Pegboard Test (PPBT) was administered to measure the movements of a person's fingers, hands, and arms [35-39]. The PPBT was initially developed by Joseph Tiffin in 1948 to test the manual dexterity of those seeking employment in industrial jobs, such as factory workers on an assembly line. Although most individuals no longer have occupations akin to factory workers, technological advancements have created new requirements for high dexterity, such as typing on a computer keyboard or messaging on a cell phone. Despite the cultural shifts in the past few decades to include technology such as mobile devices, the PPBT has been shown to be valid and reliable for wrist and hand disorders and has since been adapted for use in testing dexterity in the clinical setting [40,41].

The PPBT consists of 3 tests at 30-second intervals using the right hand, left hand, and both hands. In each test, participants were asked to pick up pins, collars, or washers from the top of the pegboard and drop them in the peg holes as rapidly as possible in 30 seconds. The score for each test was based on the total number of pins, collars, or washers that dropped in the holes correctly. A composite score was calculated by summing the scores from these 3 tests, yielding the *right+left+both hands* score. This score represents participants' overall dexterity levels. Lower *right+left+both hands* scores indicate a higher degree of dexterity impairment. On the basis of their *right+left+both hands* scores, participants were categorized into the following 3 groups reflecting their dexterity levels:

- Group 1: users with mild dexterity issues as defined by PPBT scores for the *right+left+both hands* scores ranging from 1 SD to 3 SD below the generic mean of factory workers.
- Group 2: users with moderate dexterity issues as defined by *right+left+both hands* scores >3 SD below the generic mean of factory workers.
- Group 3: users with severe dexterity issues as defined by the inability to perform the PPBT (*right+left+both hands* score=0).

Efficiency and Effectiveness

Participants were then asked to perform a set of tasks using both the original and redesigned MyMeds and SkinCare modules. The *think aloud* method for product design and development was used to gain comprehensive knowledge of participants' experiences, including any experienced frustration [42]. Specifically, participants were asked to verbally describe their intentions and actions to the researcher as they performed the following tasks:

- Task 1: schedule a new medication—participants were asked to locate the correct medication, add information about their regimen, and set up a reminder.
- Task 2: modify a medication reminder—participants were asked to change the alert time for medication.
- Task 3: respond to a medication alert—participants were asked to indicate whether a medication was taken.
- Task 4: set up a schedule to check the skin—participants were asked to set a daily alert to conduct a skin evaluation.
- Task 5: modify an alert for skin check—participants were asked to change the alert time for the scheduled skin evaluation.
- Task 6: report a skin issue—participants were asked to identify a skin issue, and then take a picture and enter data into predefined fields within the module, describing the affected skin region, including location, color, size, depth, and tissue condition.

- Task 7: update or track changes in an existing skin issue—participants were asked to reassess previously identified skin issues and track changes by taking pictures and filling out a form describing the affected skin region, including location, color, size, depth, and tissue condition.
- Task 8: set personalized configurations for user interface presentations—participants were asked to record a preferred module list, background, text size, and target size to optimize interactions. This task was conducted only for the redesigned module.

Task 8 was performed only when a participant was testing the redesigned modules.

The following variables were collected:

- Efficiency
  - Average task time: the time in seconds for a participant to complete each task was measured and then averaged across all 8 tasks.
- Effectiveness
  - Number of steps in each task: number of actions taken by the participant to complete a given task.
  - Number of mistakes in each task: when a participant reported a problem finishing a task, it was counted as a mistake. Mistakes were tallied to each task.
  - Error rate: the sum of mistakes divided by the total number of steps required to complete a task.
  - Mistake recovery: ability of participants to correct mistakes. Step-by-step observation notes were used to record the status of mistake recoveries, which were used to describe the DP experienced by a participant during mistake recovery. The DP score was calculated as the sum of weighted scores, where a lower DP score indicated better and easier performance on the task.
    1. The participant solved the problem without any help.
    2. The participant needed help solving the problem, addressed in one sentence.
    3. The participant needed help solving the problem, addressed in 2-4 sentences.
    4. The participant did not solve the problem.

Learnability and Satisfaction

Overview

Usability was measured using TUQ (Table 1). The TUQ measures constructs of usability, such as learnability and satisfaction. Learnability, as defined by Nielsen [30], assesses how easily users can accomplish a task the first time they encounter the interface and how many repetitions it takes for them to become efficient at that task. The TUQ has been shown to have high validity, reliability, and internal consistency [23]. It provides a more comprehensive evaluation of telehealth tools, given that it has combined existing sources in telemedicine (such as the Telemedicine Satisfaction Questionnaire) and computer software interface (such as the Technology Acceptance Model and the IBM Post Study System Usability Questionnaire). Participants were asked to rate the extent to which they agreed with 21 statements using a scale from 1 to 7 (minimum score 21; maximum score 147). Statements are grouped into six domains: usefulness, ease of use and learnability, interface quality, interaction quality, reliability, and satisfaction and future use. The average TUQ scores were calculated for each of the 6 domains and overall. A higher score reflects higher usability.

Table 1. Telehealth Usability Questionnaire items.

Components	Questionnaire items
Usefulness	
1	Telehealth improves my access to health care services
2	Telehealth saves me time traveling to a hospital or specialist clinic
3	Telehealth provides for my health care needs
Ease of use and learnability	
1	It was simple to use this system
2	It was easy to learn to use the system
3	I believe I could become productive quickly using this system
Interface quality	
1	The way I interact with this system is pleasant
2	I like using the system
3	The system is simple and easy to understand
4	This system is able to do everything I would want it to be able to do
Interaction quality	
1	I could easily talk to the clinician using the telehealth system
2	I could hear the clinician clearly using the telehealth system
3	I felt I was able to express myself effectively
4	Using the telehealth system, I could see the clinician as well as if we met in person
Reliability	
1	I think the visits provided over the telehealth system are the same as in-person visits
2	Whenever I made a mistake using the system, I could recover easily and quickly
3	The system gave error messages that clearly told me how to fix problems
Satisfaction and future use	
1	I feel comfortable communicating with the clinician using the telehealth system
2	Telehealth is an acceptable way to receive health care services
3	I would use telehealth services again
4	Overall, I am satisfied with this telehealth system

User Preferences

We measured user preferences by asking each participant whether they preferred the original or redesigned modules and the reasons for those preferences.

Accessibility

The following 10 accessibility features were demonstrated to participants in the redesigned app as part of the training during the study:

1. Customized module list: this feature provides the user with the ability to customize their app by hiding or showing a selected module from the home screen. The participants were able to personalize their home screens with the modules that were most applicable to them.
2. Customized text display: this feature allows the user to set up a reading size that is comfortable for them in the redesigned modules. The size, color, bold, and italic versions of titles, text, attention text, and warning text were predefined in the iMHere modules relative to the settings of the display text.
3. Customized theme: this feature allows the user to select their preferred background and text color.
4. Customized button size: customized button size was created after a user pressed their index finger on the screen to record her or her fingertip size. The smartphone then adapts button or icon size accordingly for all iMHere modules. Given the dexterity impairment in the study population, this feature improved the accuracy in making selections using a customized button target size.

5. Customized keyboard: the iMHere app provided a customized keyboard with softer keys, larger key sizes, and preconfigured characters. Customized keyboards were used primarily for the MyMeds module, where users could easily enter medication dosage information. When using the customized keypad to enter *2 tablets*, for instance, only 2 touches were needed, *2* and *tablet*. This customized keypad was designed to reduce the number of touches required on the smartphone screen.
6. Ability to take pictures of a pill or bottle: using this feature, users could take a photo of a pill or medication bottle and upload it into his or her medication schedule.
7. Color-coding: this feature matched the color with the module name. For instance, the title for the SkinCare module on the home page was highlighted in red. When navigating through the SkinCare modules, all screens under the module had a red bar.
8. Navigational short cut: this feature allowed users to create personalized settings for the home screen, such as a list of modules.
9. Text guidance: the modules provided short text cues with self-training instructional notes on the screen and were highlighted in a particular color.
10. Voice guidance: the modules used text-to-speech technology, which allowed users to listen to text guidance as audio output.

We asked participants to rank the importance of each accessibility feature, using a 10-point Likert scale (1 indicated that this feature was the most important and 10 indicated that this feature was the least important). The average ranking was then calculated for each accessibility feature.

Statistical Analysis

Descriptive statistics were calculated for the demographic and usability variables, including PPBT scores.

The α level was set at .05. All statistical analyses regarding hypotheses 1 and 2 were carried out using Wilcoxon signed-rank tests. To test the first hypothesis, the original and redesigned modules were compared in terms of efficiency (average task time) and effectiveness (number of steps, number of mistakes, error rate, and mistake recovery). As some experienced users were recruited, a secondary analysis using the Mann-Whitney *U* test was used to explore whether differences in average task time for the original and redesigned modules between experienced and inexperienced users could be because of a learning effect not mitigated by the washout period. To test the second hypothesis, the original and redesigned modules were compared in terms of usability (average overall TUQ and average TUQ domain scores).

Results

Overview

A total of 28 participants were recruited for this study; 2 (7%) participants were excluded based on the exclusion criteria: 1 (4%) user was blind, and 1 (4%) user had both vision and dexterity impairments that precluded the use of a smartphone. Moreover, 4% (1/28) of participants was not able to complete the entire protocol because of severe dexterity impairments, as assessed by PPBT scores. In addition, 4% (1/28) of participants dropped out because of scheduling conflicts. Therefore, in total, 24 participants (n=8, 33% females and n=16, 67% males) completed the entire study protocol.

Demographics and Dexterity

The demographics of the participants are presented in Table 2. Participants' ages ranged from 18 to 64 years, with an average age of 28 years (SD 6.3 years). Of the 24 participants, 14 (58%) patients had SB, 5 (21%) had SCI, 3 (13%) had cerebral palsy, 1 (4%) had muscular dystrophy, and 1 (4%) had cerebellar ataxia. In total, of the 24 participants, 22 (92%) patients were right-hand dominant, 21 (88%) were smartphone users, 2 (8%) were regular mobile phone users, and 1 (4%) participant did not use any mobile device; 12 (50%) participants had used a mobile phone for <2 years, and 20 (83%) participants used a smartphone for at least 60 minutes per day. In addition, 21% (5/24) of participants had finished graduate-level education, while 71% (17/24) of participants had received a high school or equivalent education.

Table 2. Participant demographics (N=24).

Demographic details	Values
Age (years), mean (SD)	28 (6.3)
Gender, n (%)	
Male	15 (63)
Female	9 (38)
Highest level of education, n (%)	
High school	17 (71)
Higher education	5 (21)
Disability, n (%)	
Spina bifida	14 (58)
Spinal cord injury	5 (21)
Cerebral palsy	3 (13)
Muscular dystrophy	1 (4)
Cerebellar ataxia	1 (4)
Type of phone, n (%)	
Regular	2 (8)
Smart	21 (88)
N/A[a]	1 (4)
Years of use, n (%)	
<2	12 (50)
>2	11 (46)
N/A[a]	1 (4)
Daily use, n (%)	
<60 min/day	3 (13)
>60 min/day	20 (83)
N/A[a]	1 (4)

[a]N/A: not applicable.

Of the 24 participants, 7 (29%) participants had previously used the iMHere modules (*experienced*), and 17 (71%) participants had not previously used any iMHere modules (*inexperienced*). The experienced participants had stopped using iMHere for at least 4 months before participating in this study, a washout period that we expected the participants did not carryover significant learning from previous experience. Of the 7 experienced users, 4 (57%) participants remembered approximately 5% of the process to complete the tasks in the original modules and approximately 10% of the process in the redesigned modules. Furthermore, 43% (3/7) of participants had no recollection of how to use the modules.

All participants' PPBT scores (*right+left+both hands*) were at least 1 SD below the generic mean (46.8, SD 4) of factory workers (Multimedia Appendix 1). There were 8 participants in group 1, 12 participants in group 2, and 5 participants in group 3.

Efficiency: Average Task Time

Table 3 shows the average time in seconds for all participants to complete tasks 1-7 using the original and redesigned modules. The average time for the 24 participants to complete tasks 1-7 in the original modules was approximately 48 seconds. This time dropped by 35% to 31 seconds when completing the tasks using the redesigned modules. Participants' speed in completing tasks 1, 2, 4, and 6 improved by >30% when comparing the redesigned modules with the original modules. A significant difference was found in the average task time for all tasks, except task 3, when comparing the original with the redesigned modules. Overall, a Wilcoxon signed-rank test showed that the total average task time for each participant was significantly different between the original and the redesigned modules (W=0.0; Z=−4.3; P<.001).

Table 3. Comparison of the average task time for all participants.

Tasks	Original modules (task time in seconds), mean (SD)	Redesigned modules (task time in seconds), mean (SD)	Time difference, seconds (%)	Statistics		
				W value	*Z* value	*P* value
Task 1: schedule a medication alert	110.5 (36.5)	68.9 (23.1)	−41.7 (−37.7)	3	−4.2	<.001
Task 2: modify a medication alert	39.6 (15.2)	25.1 (11.1)	−14.5 (−36.5)	24	−3.6	<.001
Task 3: respond to a medication alert	4.2 (3.1)	4.3 (2.9)	0.1 (1.8)	144	−0.2	.85
Task 4: schedule skin check	25.3 (11.2)	16.7 (6.6)	−8.5 (−33.7)	17	−3.8	<.001
Task 5: modify a skincare alert	21.8 (9.4)	16.5 (9.5)	−5.3 (−24.4)	56	–2.7	.007
Task 6: report a new skin problem	81.2 (17.8)	48.5 (12.0)	−32.7 (−40.2)	1	−4.3	<.001
Task 7: track the changes of a skin issue	56.0 (15.2)	38.8 (11.0)	−17.2 (−30.6)	9	−4.0	<.001

The average time in seconds to complete tasks using the original and the redesigned modules for the 29% (7/24) experienced participants and the 71% (17/24) inexperienced participants is shown in Table 4. A secondary analysis revealed no significant difference in average task time between the experienced (n=7; mean 49.0, SD 36.6) and inexperienced participants (n=17; mean 48.0, SD 37.4) when using the original modules (U=59; Z=−0.03; P=.98), or between the experienced (n=7; mean 31.6, SD 23.8) and inexperienced participants (n=17; mean 31.1, SD 21.7) when using the redesigned modules (U=59, Z=−0.03; P=.98).

Table 4. Experienced versus inexperienced: average task time for all participants.

Tasks	Original modules		Redesigned modules	
	Experienced (task time in seconds), mean (SD)	Inexperienced (task time in seconds), mean (SD)	Experienced (task time in seconds), mean (SD)	Inexperienced (task time in seconds), mean (SD)
Task 1: schedule a medication alert	109.0 (49.2)	111.2 (31.8)	74.1 (36.5)	66.7 (15.8)
Task 2: modify a medication alert	46.0 (19.5)	37.0 (12.9)	21.4 (7.7)	26.7 (12.1)
Task 3: respond to a medication alert	4.2 (1.6)	4.2 (3.6)	3.9 (1.2)	4.5 (3.3)
Task 4: schedule a skin check	25.4 (15.8)	25.2 (9.4)	16.5 (7.1)	16.8 (6.6)
Task 5: modify a skincare alert	21.2 (10.1)	22.0 (9.5)	18.4 (13.1)	15.7 (8.0)
Task 6: report a new skin problem	81.1 (18.4)	81.2 (18.1)	47.6 (12.7)	48.9 (12.0)
Task 7: track the changes in skin issues	58.7 (18.4)	54.9 (17.1)	39.2 (9.2)	38.7 (12.0)

As shown in Table 5, participants with severe dexterity issues (group 3) required approximately 55 seconds on average to complete the tasks using the original modules. The time to complete the tasks improved by 40% (33 seconds) using the redesigned modules, which was the largest improvement among the 3 groups. The speed of participants with mild and moderate dexterity impairments (groups 1 and 2) to complete these tasks with the redesigned modules improved by >30%.

Table 5. Group comparison of the average task time.

Tasks	Original modules (task time in seconds), mean (SD)	Redesigned modules (task time in seconds), mean (SD)	Time difference, seconds (%)
Group 1	44.6 (8.0)	31.7 (5.7)	−12.8 (−28.8)
Group 2	47.9 (11.4)	30.2 (6.5)	−17.7 (−37)
Group 3	54.9 (14.1)	35.8 (10.8)	−19.1 (−34.8)

The activities in task 8 for configuring personalized settings include choosing preferred modules, changing the background and text color, changing the text display size, and choosing the button or target size. Participants required approximately 36 seconds (SD 9.0 seconds) to complete this task. Specifically, participants with mild dexterity issues (group 1) spent 32.8 seconds (SD 7.07 seconds), participants with moderate dexterity issues (group 2) spent 34.4 seconds (SD 9.98 seconds), and those with severe dexterity issues (group 3) spent 42.2 seconds (SD 6.67 seconds) to complete this task.

Effectiveness

Overview

Table 6 shows the total number of steps to complete the tasks, the total number of mistakes committed, the calculated error rate, and the total DP scores recorded for participants completing tasks 1-7 using the original and redesigned modules.

Table 6. Comparison of total steps, mistakes, and error rate.

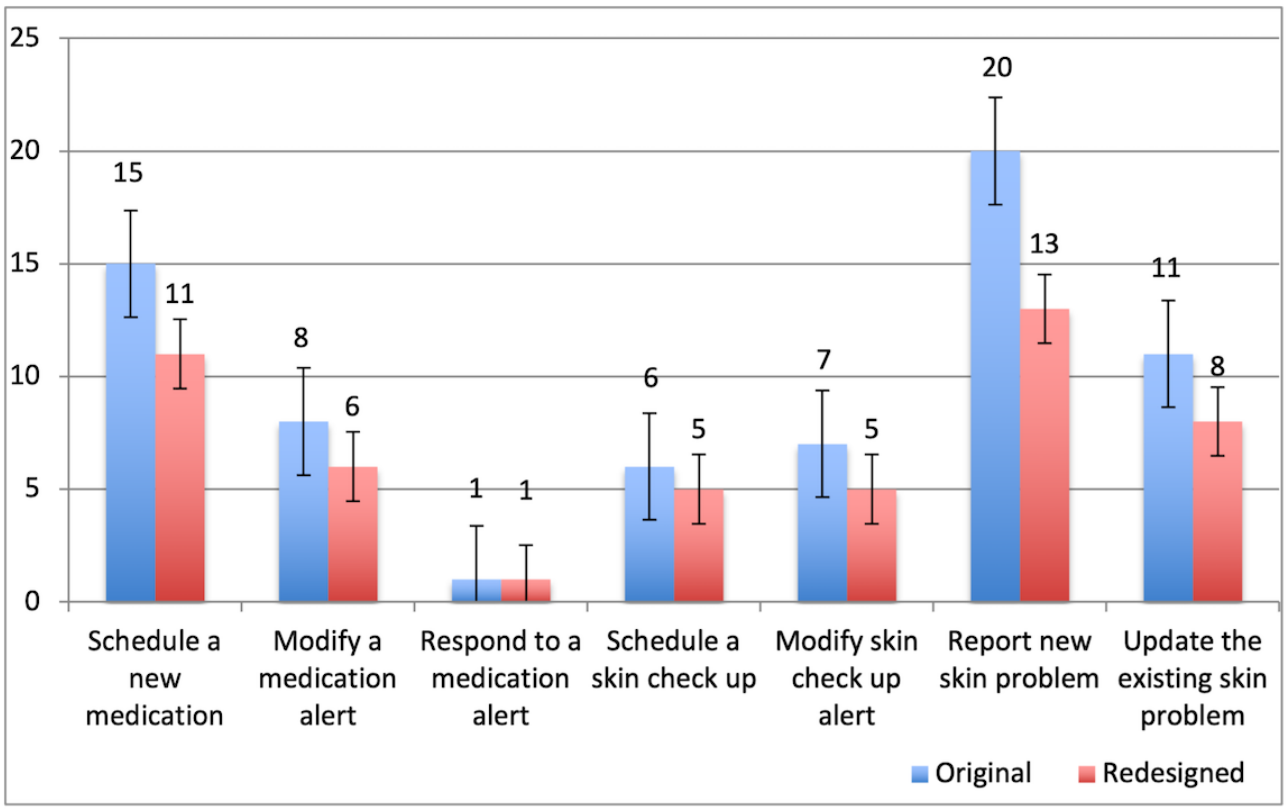
Tasks	Original modules				Redesigned modules			
	Total steps, n	Total mistakes, n	Error rate, %	Total DP[a]	Total steps, n	Total mistakes, n	Error rate, %	Total DP
Task 1: schedule a new medication	360	32	9.3	69	264	4	1.5	8
Task 2: modify a medication alert	192	21	10.9	41	144	2	1.4	4
Task 3: respond to a medication alert	24	0	0	0	24	0	0	0
Task 4: schedule a skin check	144	5	2.9	9	120	0	0	0
Task 5: modify skin check alert	168	6	3.1	12	120	3	2.5	5
Task 6: report new skin problem	480	13	2.6	21	312	4	1.3	8
Task 7: update the existing skin problem	264	16	5.9	36	192	3	1.6	5
Total	1632	93	5.7	188	1176	16	1.4	30

[a]DP: difficulty-on-performance.

Number of Steps

Figure 2 shows the average number of steps required by each participant to complete tasks 1-7 when using both the original and redesigned modules. On average, 68 steps (15+8+1+6+7+20+11) were required for a participant to complete tasks 1-7 using the original modules. This number dropped by approximately 25% to 49 steps (11+6+1+5+5+13+8) for the redesigned modules. In both modules, tasks 1 and 6 required the greatest number of steps to complete the task. A statistically significant difference was found in the number of steps for a participant to complete tasks in the original (mean 9.71, SD 6.26), and redesigned modules (W=0.0; Z=−2.2; P=.03).

Figure 2. Number of steps for participants to complete tasks.



Number of Mistakes and Error Rate

A total of 93 mistakes were identified when the participants completed the tasks using the original modules. Only 16 mistakes were identified when participants completed tasks using the redesigned modules, with an 82.8% drop rate. The reduction in the total number of mistakes for participants completing tasks 1-7 using the redesigned modules (mean 0.63, SD 1.13) compared with the original modules (mean 3.88, SD 2.66) was significantly lower (W=0.0; Z=−2.2; P=.03).

Mistake Recovery

The total DP score for participants to complete tasks 1-7 using the redesigned modules (mean 4.29, SD 3.30) was significantly lower than that for the original modules (mean 26.86, SD 23.65; W=0.0; Z=−2.2; P=.03).

While using the original module, participants were able to self-correct 22% (21/93) of the mistakes identified without any assistance (DP=1), 55% (52/93) after 1 sentence of assistance (DP=2), and 21% (20/93) after 2 sentences of assistance (DP=3). With the redesigned module, participants were able to self-correct 18% (3/16) of the mistakes without any assistance (DP=1), 73% (11/16) of the mistakes after 1 sentence of assistance (DP=2), and 6% (1/16) of the mistakes after 2 sentences of assistance (DP=3).

Learnability and Satisfaction

Figure 3 shows a comparison of the mean TUQ scores from the domain of the TUQ for the original and redesigned modules. On average, participants' usability scores improved from 83% (5.86/7, SD 0.97 points) for the original modules to 92% (6.46/7 points, SD 0.53 points) for the redesigned modules, an 8.6% improvement rate. The greatest improvements in user satisfaction were noted for *ease of use and learning* (15.45%), *interface quality* (10.97%), *interaction* (10.24%), and *reliability* (13.78%). The average TUQ scores for *usefulness*, and *satisfaction and future use* increased by >7%. The difference in usability between the original and redesigned modules was significant (W=210; Z=3.9; P<.001).

Figure 4 illustrates the average overall TUQ scores for each of the 24 participants using the original and redesigned modules. With the exception of participants 15 and 21, who had the same average overall TUQ score for both modules, all other participants had higher scores for the redesigned modules. The average overall TUQ scores were significantly different when comparing scores for the original and redesigned modules (P<.001).

Figure 3. Comparison of Telehealth Usability Questionnaire (TUQ) factors and scores.

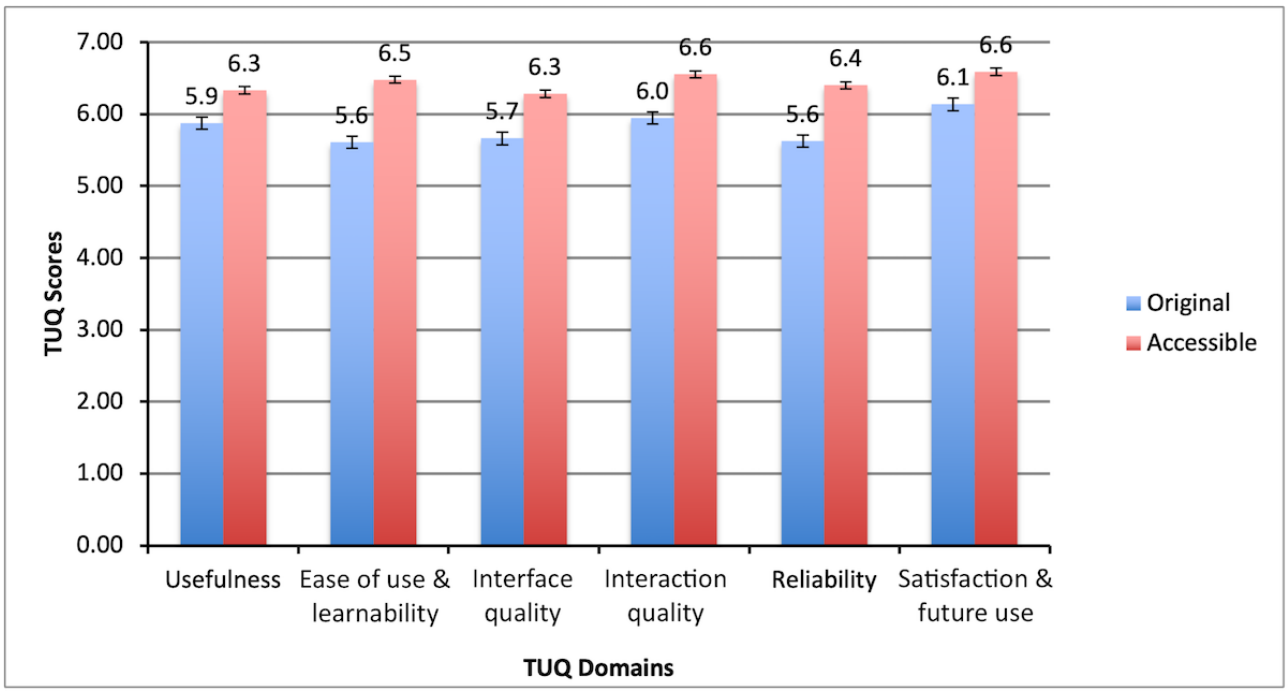
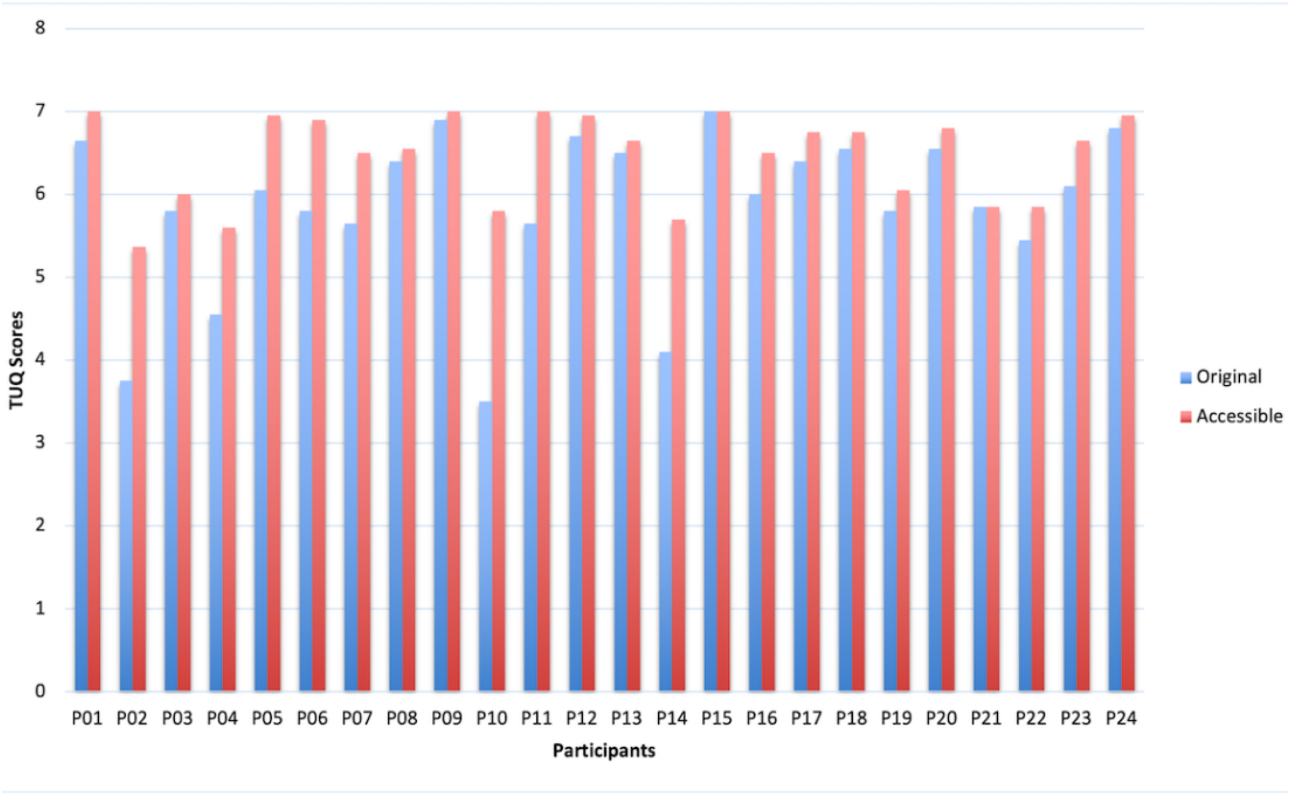


Figure 4. Telehealth Usability Questionnaire (TUQ) scores from participants.



User Preferences

Of the 24 participants, 11 (46%) tested the original modules first, followed by a test of the redesigned modules. A total of 50% (12/24) of participants tested the redesigned modules first, followed by a test of the original modules. When we asked participants' preferences regarding the use of the original or redesigned modules, 79% (19/24) of participants indicated that they preferred the redesigned modules, 13% (3/24) possibly preferred the redesigned modules, and 4% (1/24) preferred the original modules.

Participants who preferred the redesigned modules appreciated the ease of navigation and display of the redesigned modules owing to *less typing* and *larger target*. Others found the voice guidance to be useful, stating the guide *gets user's attention for directional notes*.

Only 4% (1/24) of participants preferred using the original modules, stating that it *looks clean* compared with the redesigned modules. This participant chose the picture of bamboo as a background in the redesigned modules, which made the redesigned modules look *busy*. However, the participant preferred the redesigned module in terms of flow in use compared with the original modules.

Importance of Accessibility Features

Table 7 shows rankings of the 10 new accessibility features.

Table 7. Importance of accessibility features.

Serial no	10-item Likert scale (1=most important; 10=not important)	Average scores	Ranking based on the average scores
1	Customized module list	2.8	2
2	Customized text display	4.0	9
3	Customized theme	5.3	10
4	Customized button size	3.1	3
5	Customized keyboard	3.3	4
6	Ability to take a picture of a pill or a bottle	3.8	8
7	Color-coding	3.8	7
8	Navigational short cuts	3.5	6
9	Text guidance	2.7	1
10	Voice guidance	3.3	4

Table 8 summarizes the accessibility importance rankings grouped by dexterity levels. Regardless of their dexterity level, all participants preferred using text guidance, ranking it highly across groups. Participants with mild to moderate dexterity

impairments preferred to use both voice guidance and text guidance equally. However, users with severe dexterity impairments ranked the voice guidance feature as less important. Owing to their physical limitations with respect to holding a smartphone and accessing the volume control button, participants with severe dexterity impairments had problems turning off the voice using the volume control button. The ability to change the button size and use the customized keypad was more essential for participants with severe dexterity issues.

Table 8. User preference for new accessibility features.

Features	Average scores			Ranking based on the average scores		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
Customized module list	3.7	2.2	3.2	7	1	4
Customized text display	4.0	3.8	4.4	8	7	8
Customized theme	7.3	4.5	4.2	10	10	7
Customized button size	3.6	3.0	2.8	4	4	1
Customized keyboard	3.0	3.8	3.0	2	7	2
Ability to take a picture of a pill or a bottle	4.7	3.6	3.0	9	6	2
Color-coding	3.4	3.9	3.8	3	9	6
Navigational short cuts	3.6	3.1	4.4	4	5	8
Text guidance	2.0	2.5	3.2	1	2	4
Voice guidance	3.6	2.5	5.0	4	2	10

Discussion

Principal Findings

The use of mHealth as a self-management intervention is a new field of research. The iMHere system is unique in that it is specifically designed to support the self-management of people with disabilities. A previous systematic review by Nussbaum et al [43] identified several mHealth apps relevant to the field of rehabilitation medicine and identified only 3 mHealth apps focused on self-management, including the iMHere system. The iMHere system has been shown to be feasible for use in the SB and SCI populations, and its use has been associated with improvements in self-management skills, caregiver assistance needed, frequency of urinary tract infections, and depressive symptoms [44,45]. In addition, Nguyen et al [46] used a web-based application to promote dyspnea self-management in persons with chronic obstructive pulmonary disease. Duggan et al [5] evaluated the SMART2 app in the self-management of chronic pain. Both mHealth apps demonstrated positive outcomes and effectiveness in self-management of the respective conditions they evaluated. However, there remains a paucity of apps focused on self-management in people with disabilities with motor, cognitive, and sensory impairments.

This study further adds to the literature on the usability of mHealth systems in people with disabilities with various dexterity-limiting disabilities, as it demonstrates that mHealth systems can be made more usable by improving efficiency, effectiveness, learnability, and user satisfaction.

Our first hypothesis addressed the usability constructs of efficiency and effectiveness (including errors). The efficiency and effectiveness of the redesigned modules were significantly better than those of the original modules, resulting in improved user performance and reduced user error. These changes were likely because of the design criteria that were implemented after careful consideration of how dexterity affects workflow and recovery from errors. The most apparent improvements in efficiency were seen in those with severe dexterity issues who benefited from text cues and color-coding of modules. These features allowed users to troubleshoot their own actions and reduce the overall error rate. Those with mild to moderate dexterity impairments benefited most from voice guidance, changes to button size, and custom keyboard options. Voice guidance, similar to text cues, also helped participants troubleshoot and reduce errors. The ability to change the target button size helped improve the user's accuracy. The customized keyboard simplified the process of data entry. It is important to note that the improvements in efficiency gained from these features may be a result of the modules becoming more intuitive from a cognitive perspective.

Our second hypothesis addressed learnability and satisfaction. The improved usability of the redesigned modules was also evidenced by the participants' preference for the redesigned modules. With the addition of accessibility features, we were able to further improve learnability through features such as navigational shortcuts and voice or text guidance. In addition, we added features to improve customizability, such as custom themes and lists. As seen with improvements in TUQ scores, the participants were more satisfied with the redesigned modules and would use the iMHere modules in the future. Of note, significant improvement in usability detected in the redesigned modules compared with the original modules may have been even larger because there was no ceiling effect in TUQ.

Future work on the translation of mHealth to various models of care for people with disabilities is planned. We are currently carrying out a clinical trial evaluating the community integration of people with disabilities using mHealth to supplement services provided by a community-based organization that supports

independent living. We are also carrying out implementation studies to evaluate how iMHere 2.0 can be used to deliver support to caregivers of people with disabilities and those with chronic conditions and to help facilitate long-term services and support such as caregiving services.

Study Limitations

Some limitations of this study deserve further discussion. First, we recruited a small sample, which limits the types of statistical analyses that can be performed. Second, although we redesigned all iMHere modules, this study assessed the design of only 2 modules. We chose these 2 modules because they are the most complex, containing both advanced features and basic features that are also found in the other 3 iMHere modules. As the 3 less-complex modules contain features that are similar to those tested in the more complex modules, we expect that the usability testing results for those modules would have been similar. Third, a variety of tools exist to test dexterity and usability measures. We chose the tests and measures intentionally based on the proposed usability theory but certainly, other theories, constructs, and tools are available. For instance, we did not test memorability as a measure of usability. We plan to incorporate this attribute of usability in future studies. Fourth, iMHere was not designed to support every disability or medical need, but its design is a result of research involving over 200 people with various disabilities and chronic conditions, children to older adults, and a diverse group of professionals and support personnel involved in the care of people with disabilities and chronic conditions. Finally, the participants in the study had a variety of diagnoses that resulted not only in dexterity impairments but also sensory and cognitive impairments. Thus, we were not able to determine which types of usability or accessibility issues were related to impairments other than those related to dexterity. Future studies will expand the participant population and stratify the results to further investigate the usability and accessibility needs of individuals based on their unique impairments and abilities.

Conclusions

This study demonstrated that the iMHere mHealth system became more usable for individuals with disabilities after redesigning it according to user-centered principles. Our findings demonstrate that users became more efficient and effective when using the redesigned modules. In addition, we found that the redesigned modules were easier to use and learn for the first-time users, and users were satisfied with the redesigned modules. By including the user in the iterative process to test usability, we were able to identify features in our original module that benefited from redesign. Since the publication of this work, iMHere has launched a subsequent version (iMHere 2.0) with additional features that are focused on enhancing user experience. The associated app now integrates the family and formal caregiver interface with the client app. In addition to the existing modules, additional modules focused on physical activity, nutrition, goal setting, and education were added to the app. In the future, we hope to complete usability testing with studies that incorporate memorability into user testing. With successful implementation of iMHere among our test participants, we hope to make this app available to different disability populations in the community to promote independence of self-management with improved clinical integration to bolster continuity of care.

Conflicts of Interest

BED, AF, BP, and GP are inventors of the iMHere system with no other financial interests in this technology.

Multimedia Appendix 1
Purdue Pegboard Test results (time to complete in seconds for right, left, and both hands).
[DOCX File , 18 KB - humanfactors_v9i1e23794_app1.docx]

References

1. Kane S, Jayant C, Wobbrock J, Ladner R. Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabilities. In: Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility. 2009 Presented at: Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility; Oct 25 - 28, 2009; Pittsburgh Pennsylvania USA. [doi: 10.1145/1639642.1639663]
2. Dawe M. Desperately seeking simplicity: how young adults with cognitive disabilities and their families adopt assistive technologies. In: Proceedings of the CHI 2006 Conference on Human Factors in Computing Systems. 2006 Presented at: CHI06: CHI 2006 Conference on Human Factors in Computing Systems; Apr 22 - 27, 2006; Montréal Québec Canada. [doi: 10.1145/1124772.1124943]
3. Tirado M. Role of mobile health in the care of culturally and linguistically diverse US populations. Perspect Health Inf Manag 2011 Jan 01;8:1e [FREE Full text] [Medline: 21307988]
4. Hebden L, Balestracci K, McGeechan K, Denney-Wilson E, Harris M, Bauman A, et al. 'TXT2BFiT' a mobile phone-based healthy lifestyle program for preventing unhealthy weight gain in young adults: study protocol for a randomized controlled trial. Trials 2013 Mar 18;14:75 [FREE Full text] [doi: 10.1186/1745-6215-14-75] [Medline: 23506013]
5. Duggan GB, Keogh E, Mountain GA, McCullagh P, Leake J, Eccleston C. Qualitative evaluation of the SMART2 self-management system for people in chronic pain. Disabil Rehabil Assist Technol 2015 Jan;10(1):53-60. [doi: 10.3109/17483107.2013.845696] [Medline: 24112276]

6. Kristjánsdóttir OB, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist-feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain: randomized controlled trial. J Med Internet Res 2013 Jan 07;15(1):e5 [FREE Full text] [doi: 10.2196/jmir.2249] [Medline: 23291270]
7. Disability and health. World Health Organization. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/disability-and-health [accessed 2021-11-21]
8. Wagner EH, Bennett SM, Austin BT, Greene SM, Schaefer JK, Vonkorff M. Finding common ground: patient-centeredness and evidence-based chronic illness care. J Altern Complement Med 2005;11 Suppl 1:S7-15. [doi: 10.1089/acm.2005.11.s-7] [Medline: 16332190]
9. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. JAMA 2002 Oct 16;288(15):1909-1914. [doi: 10.1001/jama.288.15.1909] [Medline: 12377092]
10. Holman H. Chronic disease--the need for a new clinical education. JAMA 2004 Sep 01;292(9):1057-1059. [doi: 10.1001/jama.292.9.1057] [Medline: 15339897]
11. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
12. 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]
13. Anderson M, Perrin A. Disabled Americans less likely to use technology. Pew Research Center. 2017. URL: https://www.benton.org/headlines/disabled-americans-are-less-likely-use-technology [accessed 2021-11-21]
14. Franz R, Wobbrock J, Cheng Y, Findlater L. Perception and adoption of mobile accessibility features by older adults experiencing ability changes. In: Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility. 2019 Presented at: ASSETS '19: The 21st International ACM SIGACCESS Conference on Computers and Accessibility; Oct 28 - 30, 2019; Pittsburgh PA USA. [doi: 10.1145/3308561.3353780]
15. Lippincot B, Thompson N, Morris J, Jones M, DeRuyter F. Survey of user needs: mobile apps for mHealth and people with disabilities. In: Computers Helping People with Special Needs. Cham: Springer; 2020.
16. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: a novel mHealth system for supporting self-care in management of complex and chronic conditions. JMIR Mhealth Uhealth 2013 Jul 11;1(2):e10 [FREE Full text] [doi: 10.2196/mhealth.2391] [Medline: 25100682]
17. Payne AY, Surikova J, Liu S, Ross H, Mechetiuc T, Nolan RP. Usability testing of an internet-based e-counseling platform for adults with chronic heart failure. JMIR Hum Factors 2015 May 08;2(1):e7 [FREE Full text] [doi: 10.2196/humanfactors.4125] [Medline: 27026267]
18. Williams PA, Furberg RD, Bagwell JE, LaBresh KA. Usability testing and adaptation of the pediatric cardiovascular risk reduction clinical decision support tool. JMIR Hum Factors 2016 Jun 21;3(1):e17 [FREE Full text] [doi: 10.2196/humanfactors.5440] [Medline: 27328761]
19. Thirumalai M, Rimmer JH, Johnson G, Wilroy J, Young H, Mehta T, et al. TEAMS (Tele-Exercise and Multiple Sclerosis), a tailored telerehabilitation mHealth app: participant-centered development and usability study. JMIR Mhealth Uhealth 2018 May 24;6(5):e10181 [FREE Full text] [doi: 10.2196/10181] [Medline: 29798832]
20. Yu DX, Parmanto B, Dicianno BE, Pramana G. Accessibility of mHealth self-care apps for individuals with Spina Bifida. Perspect Health Inf Manag 2015;12:1h [FREE Full text] [Medline: 26755902]
21. Yu DX, Parmanto B, Dicianno BE, Watzlaf VJ, Seelman KD. Accessibility needs and challenges of a mHealth system for patients with dexterity impairments. Disabil Rehabil Assist Technol 2017 Jan;12(1):56-64. [doi: 10.3109/17483107.2015.1063171] [Medline: 26153097]
22. Yu D, Parmanto B, Dicianno B. An mHealth app for users with dexterity impairments: accessibility study. JMIR Mhealth Uhealth 2019 Jan 08;7(1):e202 [FREE Full text] [doi: 10.2196/mhealth.9931] [Medline: 30622096]
23. Parmanto B, Lewis AN, Graham KM, Bertolet MH. Development of the Telehealth Usability Questionnaire (TUQ). Int J Telerehabil 2016 Jul 1;8(1):3-10 [FREE Full text] [doi: 10.5195/ijt.2016.6196] [Medline: 27563386]
24. Whitney DG, Schmidt M, Peterson MD, Haapala H. Polypharmacy among privately insured adults with cerebral palsy: a retrospective cohort study. J Manag Care Spec Pharm 2020 Sep;26(9):1153-1161. [doi: 10.18553/jmcp.2020.26.9.1153] [Medline: 32857655]
25. Cadel L, Everall AC, Packer TL, Hitzig SL, Patel T, Lofters AK, et al. Exploring the perspectives on medication self-management among persons with spinal cord injury/dysfunction and providers. Res Social Adm Pharm 2020 Dec;16(12):1775-1784. [doi: 10.1016/j.sapharm.2020.01.014] [Medline: 32057689]
26. Morley CP, Struwe S, Pratte MA, Clayton GH, Wilson PE, Dicianno BE, et al. Survey of U.S. adults with spina bifida. Disabil Health J 2020 Apr;13(2):100833 [FREE Full text] [doi: 10.1016/j.dhjo.2019.100833] [Medline: 31399347]
27. Beierwaltes P, Munoz S, Wilhelmy J. Integument: guidelines for the care of people with spina bifida. J Pediatr Rehabil Med 2020;13(4):543-548 [FREE Full text] [doi: 10.3233/PRM-200723] [Medline: 33252092]
28. Baron J, Swaine J, Presseau J, Aspinall A, Jaglal S, White B, et al. Self-management interventions to improve skin care for pressure ulcer prevention in people with spinal cord injuries: a systematic review protocol. Syst Rev 2016 Sep 06;5(1):150 [FREE Full text] [doi: 10.1186/s13643-016-0323-4] [Medline: 27600153]

29. Groah SL, Schladen M, Pineda CG, Hsieh CJ. Prevention of pressure ulcers among people with spinal cord injury: a systematic review. PM R 2015 Jun;7(6):613-636. [doi: 10.1016/j.pmrj.2014.11.014] [Medline: 25529614]
30. Nielsen J. Usability Engineering. Amsterdam: Elsevier; 1993.
31. Zeng X. Evaluation and enhancement of web content accessibility for persons with disabilities. Doctoral Dissertation, University of Pittsburgh. 2004. URL: http://d-scholarship.pitt.edu/7311/ [accessed 2021-11-21]
32. Hackett S. An exploration into two solutions to propagating web accessibility for blind computer users. Doctoral Dissertation, University of Pittsburgh. 2007. URL: http://d-scholarship.pitt.edu/10108/ [accessed 2021-11-21]
33. Scotch M, Parmanto B, Monaco V. Evaluation of SOVAT: an OLAP-GIS decision support system for community health assessment data analysis. BMC Med Inform Decis Mak 2008 Jun 09;8:22 [FREE Full text] [doi: 10.1186/1472-6947-8-22] [Medline: 18541037]
34. Saptono A. Development of an integrated telerehabilitation information management system to support remote wheelchair prescription. Doctoral Dissertation, University of Pittsburgh. 2011. URL: https://tinyurl.com/yckpm5ys [accessed 2021-11-21]
35. Desrosiers J, Hébert R, Bravo G, Dutil E. The Purdue Pegboard Test: normative data for people aged 60 and over. Disabil Rehabil 1995 Jul;17(5):217-224. [doi: 10.3109/09638289509166638] [Medline: 7626768]
36. Ozçelik IB, Purisa H, Sezer I, Mersa B, Kabakaş F, Tuncer S, et al. [Evaluation of long-term results in mutilating hand injuries]. Ulus Travma Acil Cerrahi Derg 2009 Mar;15(2):164-170 [FREE Full text] [Medline: 19353320]
37. Delp HL, Newton RA. Effects of brief cold exposure on finger dexterity and sensibility in subjects with Raynaud's phenomenon. Phys Ther 1986 Apr;66(4):503-507. [doi: 10.1093/ptj/66.4.503] [Medline: 3960975]
38. Wilson BC, Iacoviello JM, Wilson JJ, Risucci D. Purdue Pegboard performance of normal preschool children. J Clin Neuropsychol 1982 May;4(1):19-26. [Medline: 7096583]
39. Smoot B, Wong J, Cooper B, Wanek L, Topp K, Byl N, et al. Upper extremity impairments in women with or without lymphedema following breast cancer treatment. J Cancer Surviv 2010 Jun;4(2):167-178 [FREE Full text] [doi: 10.1007/s11764-010-0118-x] [Medline: 20373044]
40. Amirjani N, Ashworth NL, Olson JL, Morhart M, Chan KM. Validity and reliability of the Purdue Pegboard Test in carpal tunnel syndrome. Muscle Nerve 2011 Feb;43(2):171-177. [doi: 10.1002/mus.21856] [Medline: 21254080]
41. Buddenberg LA, Davis C. Test-retest reliability of the Purdue Pegboard Test. Am J Occup Ther 2000;54(5):555-558. [doi: 10.5014/ajot.54.5.555] [Medline: 11006818]
42. Lewis C. Using the "Thinking Aloud" Method in Cognitive Interface Design. Yorktown Heights, NY: IBM T. J. Watson Research Center; 1982.
43. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil 2019 Jan;100(1):115-127. [doi: 10.1016/j.apmr.2018.07.439] [Medline: 30171827]
44. Dicianno BE, Fairman AD, McCue M, Parmanto B, Yih E, McCoy A, et al. Feasibility of using mobile health to promote self-management in spina bifida. Am J Phys Med Rehabil 2016 Jun;95(6):425-437. [doi: 10.1097/PHM.0000000000000400] [Medline: 26488144]
45. Kryger MA, Crytzer TM, Fairman A, Quinby EJ, Karavolis M, Pramana G, et al. The effect of the interactive mobile health and rehabilitation system on health and psychosocial outcomes in spinal cord injury: randomized controlled trial. J Med Internet Res 2019 Aug 28;21(8):e14305 [FREE Full text] [doi: 10.2196/14305] [Medline: 31464189]
46. Nguyen HQ, Donesky-Cuenco D, Wolpin S, Reinke LF, Benditt JO, Paul SM, et al. Randomized controlled trial of an internet-based versus face-to-face dyspnea self-management program for patients with chronic obstructive pulmonary disease: pilot study. J Med Internet Res 2008 Apr 16;10(2):e9 [FREE Full text] [doi: 10.2196/jmir.990] [Medline: 18417444]

Abbreviations

DP: difficulty-on-performance
iMHere: Interactive Mobile Health and Rehabilitation
mHealth: mobile health
PPBT: Purdue Pegboard Test
SB: spina bifida
SCI: spinal cord injury
TUQ: Telehealth Usability Questionnaire

Edited by A Kushniruk; submitted 23.08.20; peer-reviewed by C Jones, A Miguel-Cruz; comments to author 30.09.20; revised version received 23.12.20; accepted 02.08.21; published 24.02.22.

Please cite as:
Chowdhary K, Yu DX, Pramana G, Mesoros M, Fairman A, Dicianno BE, Parmanto B
User-Centered Design to Enhance mHealth Systems for Individuals With Dexterity Impairments: Accessibility and Usability Study
JMIR Hum Factors 2022;9(1):e23794
URL: https://humanfactors.jmir.org/2022/1/e23794
doi:10.2196/23794
PMID:35200144

Original Paper

Usability Testing and Technology Acceptance of an mHealth App at the Point of Care During Simulated Pediatric In- and Out-of-Hospital Cardiopulmonary Resuscitations: Study Nested Within 2 Multicenter Randomized Controlled Trials

Johan N Siebert[1,2], MD; Laëtitia Gosetto[3*], MS; Manon Sauvage[3*], MS; Laurie Bloudeau[4], EMT-P; Laurent Suppan[2,5], MD; Frédérique Rodieux[2,6], MD; Kevin Haddad[1], RNCS; Florence Hugon[1], RNCS; Alain Gervaix[1,2], MD; Christian Lovis[2,7], MPH, MD; Christophe Combescure[2,8], PhD; Sergio Manzano[1,2], MD; Frederic Ehrler[3], PhD; PedAMINES Trial Group[9]; PedAMINES Prehospital Group[9]

[1]Department of Pediatric Emergency Medicine, Geneva Children's Hospital, Geneva University Hospitals, Geneva, Switzerland

[2]Faculty of Medicine, University of Geneva, Geneva, Switzerland

[3]Diagnostic Department, Geneva University Hospitals, Geneva, Switzerland

[4]A.C.E. Geneva Ambulances SA, Geneva, Switzerland

[5]Department of Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland

[6]Division of Clinical Pharmacology and Toxicology, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland

[7]Division of Medical Information Sciences, Department of Radiology and Medical Informatics, Geneva University Hospitals, Geneva, Switzerland

[8]Division of Clinical Epidemiology, Department of Health and Community Medicine, Geneva University Hospitals, Geneva, Switzerland

[9]See Acknowledgments

[*]these authors contributed equally

Corresponding Author:
Johan N Siebert, MD
Department of Pediatric Emergency Medicine
Geneva Children's Hospital
Geneva University Hospitals
Avenue de la Roseraie, 47
Geneva, 1205
Switzerland
Phone: 41 795534072
Email: Johan.Siebert@hcuge.ch

Abstract

Background: Mobile apps are increasingly being used in various domains of medicine. Few are evidence-based, and their benefits can only be achieved if end users intend to adopt and use them. To date, only a small fraction of mobile apps have published data on their field usability and end user acceptance results, especially in emergency medicine.

Objective: This study aims to determine the usability and acceptance of an evidence-based mobile app while safely preparing emergency drugs at the point of care during pediatric in- and out-of-hospital cardiopulmonary resuscitations by frontline caregivers.

Methods: In 2 multicenter randomized controlled parent trials conducted at 6 pediatric emergency departments from March 1 to December 31, 2017, and 14 emergency medical services from September 3, 2019, to January 21, 2020, the usability and technology acceptance of the PedAMINES (Pediatric Accurate Medication in Emergency Situations) app were evaluated among skilled pediatric emergency nurses and advanced paramedics when preparing continuous infusions of vasoactive drugs and direct intravenous emergency drugs at pediatric dosages during standardized, simulation-based, pediatric in- and out-of-hospital cardiac arrest scenarios, respectively. Usability was measured using the 10-item System Usability Scale. A 26-item technology acceptance self-administered survey (5-point Likert-type scales), adapted from the Unified Theory of Acceptance and Use of Technology model, was used to measure app acceptance and intention to use.

Results: All 100% (128/128) of nurses (crossover trial) and 49.3% (74/150) of paramedics (parallel trial) were assigned to the mobile app. Mean total scores on the System Usability Scale were excellent and reached 89.5 (SD 8.8; 95% CI 88.0-91.1) for nurses and 89.7 (SD 8.7; 95% CI 87.7-91.7) for paramedics. Acceptance of the technology was very good and rated on average >4.5/5 for 5 of the 8 independent constructs evaluated. Only the image construct scored between 3.2 and 3.5 by both participant populations.

Conclusions: The results provide evidence that dedicated mobile apps can be easy to use and highly accepted at the point of care during in- and out-of-hospital cardiopulmonary resuscitations by frontline emergency caregivers. These findings can contribute to the implementation and valorization of studies aimed at evaluating the usability and acceptance of mobile apps in the field by caregivers, even in critical situations.

Trial Registration: ClinicalTrials.gov NCT03021122; https://clinicaltrials.gov/ct2/show/NCT03021122. ClinicalTrials.gov NCT03921346; https://clinicaltrials.gov/ct2/show/NCT03921346

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-019-3726-4

(JMIR Hum Factors 2022;9(1):e35399) doi:10.2196/35399

KEYWORDS

cardiopulmonary resuscitation; drugs; emergency medical services; medication errors; mobile health; mobile apps; out-of-hospital cardiac arrest; paramedics; pediatrics; System Usability Scale; Unified Theory of Acceptance and Use of Technology; smartphone; mobile phone

Introduction

Background

Over the past few decades, health information technologies (HITs) and communication technologies have been widely adopted in health care environments to improve care provision, efficiency, quality, and patient safety while achieving cost savings [1]. Supported by the rapid spread of mobile devices and their innovative features (eg, connectivity, on-board computing capabilities, small size, and operating systems allowing mobile app development), mobile health (mHealth) has undergone considerable development to address health issues by providing medical and communication services within easy reach of end users [2-4]. As of the first quarter of 2021, approximately 5.7 million apps were available on leading web-based app stores [5]. Among them, >325,000 are mHealth apps [6]. The vast majority (65%) are wellness apps designed to be used primarily by the general public. Approximately 15% are patient-centered apps and focus on self-management of specific conditions, and the remaining 20% are medical apps intended for health care providers [7]. Unfortunately, most do not adhere to relevant medical evidence and lack expert involvement in their development or validation process through high-quality studies to support their adoption in clinical practice [2,8-13]. Even when apps are evidence-based, this does not guarantee that they will be used consistently over time. As with other HIT, their benefits can only be achieved if end users intend to adopt them [10,14].

Understanding users' HIT adoption behavior is a long-standing topic in the literature and could lead to improvements in the acceptability and use of mHealth apps. Several models have been proposed to predict and understand users' acceptance and use of HIT. Two of the most commonly used models are the Technology Acceptance Model (TAM) [15,16], or its extended versions (TAM Task–Technology Fit [TTF] [17], TAM2 [18], and TAM3 [19]), and the Unified Theory of Acceptance and Use of Technology (UTAUT; version 1 [20] or 2 [21]). However, a systematic review identified that mobile apps were the least frequently studied HIT application areas using the TAM, with only 15 studies published between 1989 (ie, the year when the TAM was introduced) and 2017 [22]. Although the UTAUT has a better explanatory power for technology acceptance [23] and has been extensively used in studies related to the adoption of HIT, user acceptance and use of medical apps from the perspective of caregivers have been little investigated by this model [24]. In addition, even if usability has been identified as a key component of good practice in the development of digital apps [25], the number of medical apps that publish their usability evaluation results remains scarce [26]. To this end, the validated System Usability Scale (SUS) has been identified as the most frequently used questionnaire [27-29]; however, it concerns mostly mobile apps designed to support patients in health self-management and not caregivers [26]. As a result, little is known about health professionals' willingness to implement and use medical apps in clinical care [26,30,31].

Previous Work

In previous randomized trials, we reported fewer medication errors and a shorter time to drug preparation and delivery during out-of-hospital cardiac arrest (OHCA) and in-hospital post–cardiac arrest scenarios when using the PedAMINES (Pediatric Accurate Medication in Emergency Situations) app than conventional preparation methods [32,33]. This evidence-based app was designed as a step-by-step guide for the preparation and delivery of intravenous drugs to address the unmet need for reducing pediatric medication errors [34]. Recent findings showed that this app was also able to reduce acute perceived stress while preparing emergency drugs in a prehospital setting during pediatric OHCA in a simulated model [35]. However, its usability and technology acceptance by frontline caregivers remains to be determined.

Aim

This study aims to investigate the usability of the PedAMINES app by both advanced paramedics with drug preparation

autonomy and nurses at the point of care to gain insight into their perceptions of its adoption as an approach to facilitate emergency drug preparation in pediatric life-threatening situations. In addition, we measure the technology acceptance of the app for its intended purpose and user satisfaction with its use. We hypothesize that this approach would help estimate the likelihood of adoption of the app for implementation among its future target users.

Methods

Study Design

This is a nested, overlapping study within the context of 2 prospective, multicenter, randomized controlled trials registered at ClinicalTrials.gov (NCT03021122 and NCT03921346). The parent trials had the broader and primary aim of assessing rates of medication dosing errors during simulation-based pediatric OHCA and in-hospital post–cardiac arrest scenarios using a high-fidelity manikin [32,33]. For that purpose, participants were randomly assigned (1:1 ratio) to prepare the drugs either with the support of the app (intervention group) or by conventional methods (control group). The trial protocols containing details of the scenarios have been previously published [36,37]. Both trials were performed in accordance with appropriate guidelines [38,39] and followed the CONSORT (Consolidated Standards of Reporting Trials) reporting guidelines [40].

Technology Acceptance Terminology and Definition

The terms *technology acceptability*, *acceptance*, and *adoption* are often used confusingly or interchangeably in the mHealth literature. For the purposes of this paper, the term *technology acceptance* is used and should be understood as referring to *initial acceptance of use*, as recently defined by Nadal et al [41], to explicitly distinguish between the different temporal stages of technological acceptance. It refers to users' first interactions with the app at the preadoption stage before sustained acceptance of its use in the postadoption stage.

Setting and Intervention

The first trial [32] was conducted at 14 emergency medical services (EMSs). An out-of-hospital child's bedroom environment was simulated at each EMS center to resemble, as much as possible, a standard environment where paramedics would have to intervene. Participants were tested on sequential preparations of four direct intravenous emergency drugs of varying degrees of preparation difficulty (epinephrine, midazolam, 10% dextrose, and sodium bicarbonate) during simulated pediatric OHCA. The second trial [33] was conducted as a crossover study at 6 pediatric emergency departments. Participants were tested on the preparation of continuous infusions (dopamine and norepinephrine) during simulated pediatric immediate post–cardiac arrest scenarios. Procedures were standardized across all sites to follow the same chronological progression and range of difficulty to ensure that each participant was exposed to exactly the same case in their respective setting, with similar challenges in app use and decision-making. The scripted and uniform delivery of the scenarios throughout the studies was strictly preserved to minimize confounders. Importantly, we did not organize pretests to minimize preparation bias so as not to influence the usability and acceptability of the app based on previous experiences. In both trials, the app was interfaced on an Apple iPad with the latest version of iOS; however, the app works identically on the Android OS (Google Inc). In both trials, all participants who had used the app were surveyed with self-administered questionnaires (refer to the following sections) immediately after completing the scenarios, with the necessary precautions taken to ensure that participants could not communicate with each other. During completion of the questionnaires, no interaction occurred between the participants and investigators other than those related to detailing an item upon the request of a participant.

Participants

All registered paramedics and pediatric nurses working at 14 EMSs and 6 academic and community Swiss pediatric emergency departments, respectively, were eligible for inclusion in the study. Participants were of both sexes, of all ages, of different levels of pediatric experience, and from different regions of Switzerland, a pluralistic country with several official languages (French, German, and Italian), to diversify participant characteristics and limit selection bias. Apart from language translation, no other transcultural adaptations were required or made to the app. The inclusion criteria were to have followed a standardized 5-minute introductory course on the use of the mobile device app and written informed consent. This introductory course was not intended at this stage to test the usability of the app but to explain its use for the upcoming intervention. Only participants who had used the app (ie, all those enrolled in the in-hospital study [crossover trial] and half of those enrolled in the out-of-hospital study [parallel group trial]) were eligible for this study. All participants were assumed to have equivalent experience and competence with intravenous drug preparation and dose calculation, as this is part of their regular practice and training background. Given that this study was nested within both randomized trials, the number of participants queried in each trial stemmed from the power calculations set for the primary outcome, with a 2-sided risk α of .05 and a power of 90%. Blinding to the purpose of each trial during recruitment was maintained to minimize the preparation bias. Participants were unblinded after randomization at the beginning of the scenario. Although the intervention could not be masked, all investigators remained unaware of the outcomes until all data were unlocked for analysis at the end of the trial.

Measurement Instruments

On the day of participation after randomized allocation and before scenarios started, each participant was required to complete a survey collecting data regarding their demographic characteristics and health care training. Five-point Likert-type scales, ranging from 1=*strongly disagree* to 5=*strongly agree*, were used to assess (1) their experience in the use of smartphones and tablets, (2) satisfaction with current supports at their disposal to prepare emergency intravenous drugs, (3) perceived mastery of the preparation of these drugs, (4) propensity to use technological tools in emergency situations, and (5) attitude toward the introduction of technological tools

to facilitate the preparation of intravenous drugs during an emergency. Each participant was then exposed to the simulated out-of-hospital or in-hospital scenarios.

After the scenarios were completed, the perceived usability of the app was measured using the reliable SUS designed by Brooke [27]. According to the International Organization for Standardization, the SUS assesses effectiveness (ie, the ability of users to use the product), efficiency (ie, the effort to use the product), and satisfaction (ie, how the users felt when using the product). It comprises a 10-item questionnaire with 5 response options for each item based on their level of agreement, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). According to the scoring system by Brooke, for odd-numbered (1, 3, 5, 7, and 9) statements (the positively worded items), the score contribution is equal to the scale position minus 1 (eg, *strongly agree*: 5 – 1=4). For even-numbered (2, 4, 6, 8, and 10) statements (the negatively worded items), the score contribution is equal to 5 minus the scale position (eg, *strongly agree*: 5 – 5=0). Each score contribution falls within the range of 0 to 4. The participants' scores for each item are then added up together and multiplied by 2.5 to convert the original scores of 0 to 40 to 0 to 100. Although the scores range from 0 to 100, these are not percentages of usability. The higher the score, the better the usability (ie, 0=*very poor perceived usability* and 100=*excellent perceived usability*). To obtain an SUS score of 100, the respondent must answer 5 to all odd questions and 0 to all even questions. The original SUS items are presented in Multimedia Appendix 1 [27]. When translating the SUS questionnaire from its original version to French, German, and Italian [42,43], we replaced the general term *system* with the specific term *PedAMINES*.

Acceptability and usability testing of the app was also assessed in both trials using a tailored 26-item technology acceptance self-administered survey. Scales and queries with high levels of internal consistency (ie, Cronbach $\alpha > .70$) were derived and slightly adapted from prior research to fit the trial's context (Multimedia Appendix 2) [44]. This integrative model gathers the 2 most commonly used models in the literature that inform technology acceptance [45], namely the TAM [15] and UTAUT [20,21], as well as additional dimensions of technology acceptance models [19,20,46-50], with the following eight core constructs: (1) perceived usefulness (4 items), (2) perceived ease of use (4 items), (3) TTF (4 items), (4) performance expectancy (3 items), (5) impact on image (2 items), (6) personal innovativeness (3 items), (7) acceptance (3 items), and (8) behavioral intention to use the technology (3 items). The items measured the constructs by asking participants to agree or disagree with statements using 5-point Likert-type scales, ranging from *strongly disagree* (1) to *strongly agree* (5). An item of the *image* construct (ie, Query 34, Image Expectancy 1; *People in my organization who use the tool have a high profile*), as well as the constructs *job security* and *facilitating conditions*, were dropped as they were irrelevant in the context under study. The final survey is depicted in the *Results* section. According to the original UTAUT model, gender, age, experience, and voluntariness of use are identified as moderators that affect technology beliefs and use [20]. Among these, we expected age to influence the attitude toward the use of the app, knowing that young adults have been more exposed to new technologies during their education and daily practice than older adults. Therefore, we selected age as a factor of interest to correlate with the technology acceptance self-administered survey items. Older participants were expected to have a greater reluctance in introducing the app into their daily practice.

Finally, a question using a 10-point Likert scale was administered to participants to measure their perceived satisfaction with the use of either the app or conventional preparation methods to prepare the drugs during the resuscitation scenarios (*on a scale of 1 to 10, how satisfied were you with your preparation experience?*). Data collection was conducted on site using Microsoft Excel spreadsheet (version 2011; Microsoft Corporation) and REDCap (Research Electronic Data Capture; Vanderbilt University) database. The investigators double-checked that the questionnaires were fully and accurately completed on site. Only study investigators had access to the data.

Statistical Analysis

Age and work experience were assessed using the Spearman correlation coefficient. The mean global SUS score was reported using SD and 2-sided 95% CIs. Items on the SUS questionnaire were described using frequencies. For the technology acceptance self-administered survey questionnaire, scores on the 8 technology acceptance dimensions were described using the means, SDs, and frequencies of participants with a score of ≥4. The items were described in a similar manner. The association between dimension scores and participant ages was examined using the Spearman correlation coefficient. Paramedics and nurses were analyzed independently, without any comparison made. Statistical tests on the correlation coefficients were 2-tailed, with a significance level of 5%. Data analysis was conducted using R for Windows (version 4.0.2; R Core Team).

Ethics Approval

Both trials received a declaration of no objection by the Geneva Cantonal Ethics Committee and Swiss Ethics as their purpose was to examine the effect of the intervention on health care providers. Both trials were registered at ClinicalTrials.gov (NCT03021122 and NCT03921346). The trials were conducted in accordance with the principles of the Declaration of Helsinki [51], Good Clinical Practice guidelines [52], the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online TeleHealth; Multimedia Appendix 3) [39], and the Reporting Guidelines for Health Care Simulation Research [38].

Results

Overview

A total of 202 participants using the app (74, 36.6%, paramedics and 128, 63.4%, nurses) completed the scenarios and questionnaires without dropouts. Table 1 summarizes their demographic and health care characteristics.

Table 1. Baseline characteristics (N=202).

Characteristics	Paramedics (n=74)	Nurses (n=128)
Age (years), mean (SD)	35.7 (7.3)	37.2 (9.7)
Age (years), n (%)		
<30	17 (23)	36 (28.1)
30-39	35 (47.3)	41 (32)
40-49	18 (24.3)	35 (27.3)
≥50	4 (5.4)	16 (12.5)
Gender, n (%)		
Female	25 (33.8)	121 (94.5)
Male	49 (66.2)	7 (5.5)
Work experience in years since certification, mean (SD)	7.8 (6)	13.6 (9.1)
Work experience in years since certification, n (%)		
<5	26 (35.1)	23 (18)
5-9	26 (35.1)	24 (18.8)
10-19	15 (20.3)	47 (36.7)
≥20	7 (9.5)	34 (26.6)
Smartphone user, n (%)		
No	0 (0)	5 (3.9)
Yes	74 (100)	123 (96.1)
Comfortable using a smartphone, n (%)		
Strongly disagree	1 (1.4)	14 (12.1)
Disagree	3 (4.1)	32 (27.6)
Neither disagree nor agree	12 (16.2)	55 (47.4)
Agree	38 (51.4)	15 (12.9)
Strongly agree	20 (27)	0 (0)
Missing data	0 (0)	12 (9.4)
Last preparation of emergency drugs (months)[a], n (%)		
Never	9 (12.2)	57 (44.5)
<6	15 (20.3)	10 (7.8)
6-12	17 (23)	13 (10.2)
12-24	15 (20.3)	12 (9.4)
>20	18 (24.3)	36 (28.1)
Satisfaction with the media currently available to prepare emergency drugs, n (%)		
Strongly disagree	9 (12.2)	9 (7.1)
Disagree	18 (24.3)	23 (18.3)
Neither disagree nor agree	25 (33.8)	63 (50)
Agree	21 (28.4)	26 (20.6)
Strongly agree	1 (1.4)	5 (4)
Missing data	0 (0)	0 (0)
Mastering the preparation of emergency drugs, n (%)		
Strongly disagree	9 (12.2)	35 (27.3)
Disagree	15 (20.3)	37 (28.9)
Neither disagree nor agree	30 (40.5)	35 (27.3)

Characteristics	Paramedics (n=74)	Nurses (n=128)
Agree	19 (25.7)	17 (13.3)
Strongly agree	1 (1.4)	4 (3.1)
Missing data	0 (0)	0 (0)
In favor of introducing technological tools to assist in emergency drug preparation, n (%)		
Strongly disagree	0 (0)	1 (0.8)
Disagree	0 (0)	0 (0)
Neither disagree nor agree	6 (8.1)	4 (3.1)
Agree	24 (32.4)	45 (35.2)
Strongly agree	44 (59.5)	78 (60.9)
Missing data	0 (0)	0 (0)

[a]For the nurses, emergency drugs meant vasoactive drugs in continuous infusion.

Although there was a wide age distribution, most respondents were aged between 30 and 49 years, with the nursing group having slightly more older participants. A strong correlation (r=0.72; P<.001) between paramedics' age and work experience, expressed in years, was found, as well as a very strong correlation (r=0.95; P<.001) for nurses.

Usability Testing

As shown in Figure 1, the mean total SUS scores for the app were 89.7 (SD 8.7; 95% CI 87.7-91.7) for paramedics and 89.5 (SD 8.8; 95% CI 88.0-91.1) for nurses' quotation, which qualifies the tool as between *excellent* and the *best imaginable*, according to Bangor et al [53]. All scores were at least >60, spanning from 62.5 (2/202, 1% of people) to 100 (26/202, 12.9% of people; Figure 2). Figure 3 provides a visual overview of the distribution of the item responses on the SUS. SUS total score was not significantly associated with participants' age (paramedics: r=−0.05 and P=.66; nurses: r=−0.01 and P=.91), which is likely related to highly skewed scores toward high usability with little variability among respondents (Multimedia Appendix 4). The other demographic moderators detailed in Table 1 did not have any significant effect on participants' acceptance of the app.

Figure 1. Overall System Usability Scale (SUS) scores to assess the usability of the PedAMINES (Pediatric Accurate Medication in Emergency Situations) app. The SUS score is located on a normalized scale ranging from a minimum score of 0 to a maximum of 100 [27]. Adjective ratings provide an interpretation of the SUS score [53]. The SUS also provides letter grades, similar to those used in the traditional school grading system [54]. The acceptability ranges indicate whether the tool is acceptable or not. Red dots represent the mean SUS score in paramedics and blue dots in nurses. Capped blue and red lines represent the 5th and 95th percentiles. Crosses represent medians (paramedics: 92.5, 5th-95th percentiles: 74.125-100; nurses: 90, 5th-95th percentiles: 72.5-100).

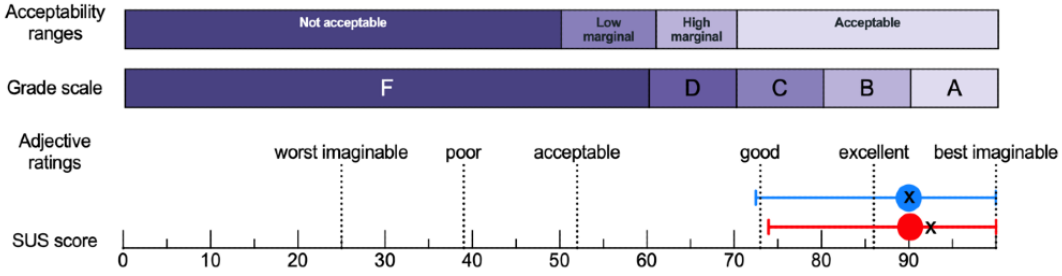


Figure 2. Distribution of counts of System Usability Scale (SUS) total scores. Red dots denote paramedics; blue dots denote nurses.

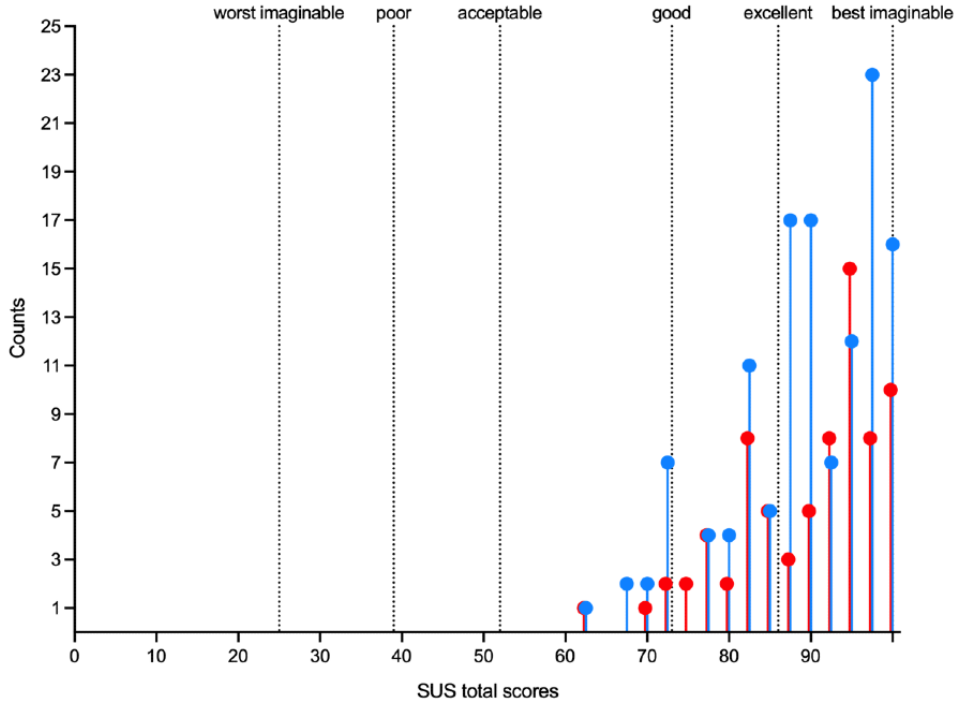
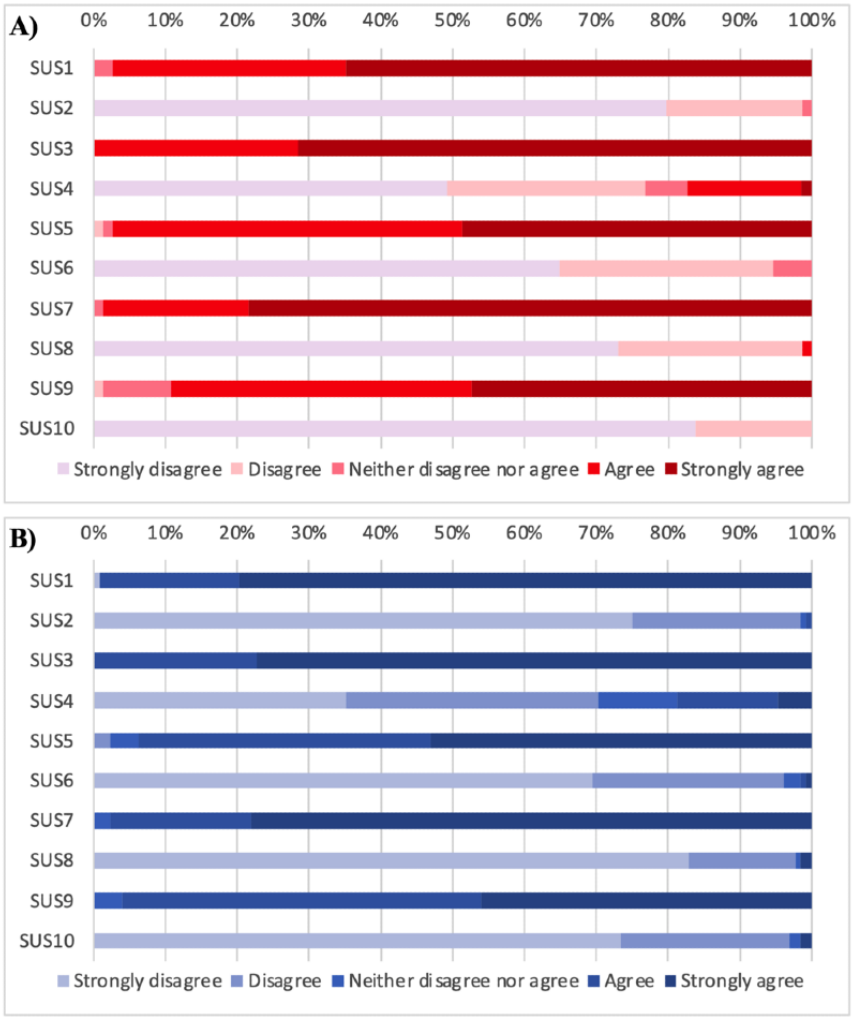


Figure 3. Percent distribution of item responses by (A) paramedics (n=74) and (B) nurses (n=128) on the (inversed) System Usability Scale (SUS) items. The SUS comprises 10 items (numbered as SUS1 to SUS10).



Technology Acceptance

Table 2 shows the overall perceptions of paramedics and nurses regarding the mobile app according to the adapted technology acceptance self-administered survey constructs. The paramedics and nurses largely agreed (mostly with scores ≥4) that the app (1) could enhance their performance during emergency drug preparation (perceived usefulness), (2) was easy to use (perceived ease of use), (3) met the requirements necessary to effortlessly address the complexity of the drug preparation task when handled by a technology tool (TTF), (4) could help them achieve performance gains during this procedure (performance expectancy), (5) elicited their intention to use it (attitude toward using a technology), (6) is acceptable for emergency drug preparation (acceptance), and (7) would be intentionally used for this purpose over a long period (intention to use). Intention to use the app was the highest-rated construct, with the item *Assuming I had access to PedAMINES, I intend to use it* receiving the highest agreement among participants (Table 2). The paramedics and nurses agreed to a lesser extent that the app would enhance the adopter's social image or status in their organization. Participants had the least agreement on the item *Having PedAMINES will be a status symbol in my organization* from the image construct. As with the SUS, the technology acceptance self-administered survey scores were skewed toward high scores of acceptance of the app and intention to use it. Only the constructs *image*, *attitude toward technology use*, and *acceptance* showed a somewhat greater spread across the score range. A weak but significantly negative correlation between nurses' age and attitude toward app use ($r=-0.27$; $P=.003$) was identified. Although the TAM used here or the original UTAUT from which it was derived can be influenced by four key moderating variables (ie, age, experience, gender, and voluntariness of use), the effect of the latter two on the acceptance constructs was not analyzed in this study, given the lack of significant variation in these moderators across individuals in the same setting [55]. Experience was highly correlated with age and, therefore, was not analyzed. No significant correlation between age and other items of the technology acceptance self-administered survey constructs was found, and no correlation was observed among paramedics. Multimedia Appendix 5 presents the results of the technology acceptance self-administered survey items by score range.

Satisfaction

The app obtained a mean overall satisfaction score of 9.1 (SD 0.9) out of 10 among paramedics and 9.4 (SD 1.0) among nurses.

Table 2. Results of the technology acceptance survey by items.

Constructs and items	Definition[a] and item wording	Paramedics (n=74)		Nurses (n=128)	
		Values, mean[b] (SD)	Score ≥4, n (%)	Values, mean (SD)	Score ≥4, n (%)
PU[c]	The degree to which an individual perceives that using the system leads to enhanced personal performance [15,18]	4.69 (0.44)	66 (89.1)	4.79 (0.37)	122 (95.3)
PU1	Using PedAMINES[d] helps me to prepare emergency drugs more quickly.	4.69 (0.62)	68 (91.9)	4.88 (0.32)	128 (100)
PU2	Using PedAMINES helps me to prepare emergency drugs better.	4.73 (0.58)	71 (95.9)	4.74 (0.52)	123 (96.1)
PU3	Using PedAMINES makes it easier for me to prepare emergency drugs.	4.69 (0.52)	72 (97.3)	4.81 (0.51)	126 (98.4)
PU4	Using PedAMINES enhances my effectiveness in drug preparation.	4.66 (0.50)	73 (99)	4.74 (0.61)	121 (94.5)
PEOU[e]	The degree to which an individual perceives that using the system will be free from physical or mental efforts [15,18,56]	4.61 (0.35)	72 (97.3)	4.76 (0.28)	126 (98.4)
PEOU1	It is easy to get PedAMINES to do what l want it to do.	4.41 (0.94)	63 (85.1)	4.78 (0.47)	125 (97.7)
PEOU2	Overall, I find PedAMINES is easy to use.	4.74 (0.47)	73 (98.6)	4.86 (0.35)	128 (100)
PEOU3	It is easy for me to become skillful in using PedAMINES.	4.89 (0.31)	74 (100)	4.88 (0.33)	128 (100)
PEOU4	I often become confused with PedAMINES' features when I used it.	4.41 (0.55)	72 (97.3)	4.53 (0.59)	124 (96.9)
TTF[f,g]	The degree to which an individual perceives that using the system fits the requirements of a particular task [17,44,48]	4.49 (0.54)	63 (85.1)	4.64 (0.43)	121 (95.3)
TTF1	PedAMINES has the functionalities l need to accomplish my tasks.	4.55 (0.62)	71 (95.9)	4.69 (0.52)	126 (98.4)
TTF2	PedAMINES' functionalities give me exactly what I need for my work.	4.47 (0.67)	67 (91)	4.61 (0.54)	125 (97.7)
TTF3	PedAMINES is very well suited to my work.	4.70 (0.54)	71 (95.9)	4.77 (0.44)	127 (99.2)
TTF4	Using PedAMINES is compatible with most aspects of my work.	4.22 (0.90)	65 (87.8)	4.50 (0.68)	116 (91.3)
PE[h]	The degree to which an individual perceives that using the system will help the user attain gains in job performance [20,44]	4.55 (0.59)	61 (82.4)	4.58 (0.57)	115 (89.8)
PE1	Using PedAMINES, I get better chances to improve my professional position.	4.28 (0.96)	59 (79.7)	4.21 (1.0)	101 (78.9)
PE2	Using PedAMINES will help me improve or continue to help to improve emergency drugs preparation.	4.61 (0.74)	67 (90.5)	4.77 (0.55)	125 (97.7)
PE3	Using PedAMINES will increase the quality of my drug preparation.	4.74 (0.57)	71 (95.9)	4.77 (0.52)	124 (96.9)
II[i,j]	The extent to which an individual perceives that using the system enhances one's image or status in ones' social system [20,50]	3.46 (0.96)	30 (40.5)	3.24 (0.98)	41 (32.5)[k]
II1	People in my practice setting who use PedAMINES will have more prestige than those who do not.	3.84 (1.06)	47 (63.5)	3.53 (1.13)	71 (55.9)[l]
II2	Using PedAMINES will be a status symbol in my practice setting.	3.08 (1.16)	22 (29.7)	2.94 (1.09)	36 (28.6)[k]

Constructs and items	Definition[a] and item wording	Paramedics (n=74)		Nurses (n=128)	
		Values, mean[b] (SD)	Score ≥4, n (%)	Values, mean (SD)	Score ≥4, n (%)
PI[m,n]	The extent to which an individual has an innate propensity (willingness) toward trying any new technology [50,57]	4.14 (0.58)	48 (64.9)	3.97 (0.77)	73 (57.5)[l]
PI1	If I heard about a new technology, I would look for ways to experiment with it.	4.55 (0.58)	71 (95.9)	4.39 (0.73)	116 (91.3)[l]
PI2	Among my peers, I am usually the first to try out new technologies.	3.59 (0.95)	42 (56.8)	3.42 (1.03)	62 (48.8)[l]
PI3	I like to experiment with new technologies.	4.28 (0.67)	65 (87.8)	4.09 (0.85)	102 (80.3)[l]
A[o,p]	The extent to which individuals accept to use a new technology [56]	4.32 (0.61)	56 (75.7)	4.30 (0.63)	100 (79.4)[k]
A1	In my opinion, it would be desirable to use PedAMINES in addition to conventional preparation methods for emergency drugs.	4.53 (0.74)	69 (93.2)	4.40 (0.90)	111 (88.1)[k]
A2	It would be good to use PedAMINES more than the conventional methods for the preparation of emergency drugs.	4.39 (0.70)	67 (90.5)	4.45 (0.78)	114 (89.8)[l]
A3	I think it would be highly desirable to use only PedAMINES instead of conventional methods for the preparation of emergency drugs.	4.04 (1.07)	54 (73)	4.06 (1.14)	94 (74)[l]
BI[q,r]	Individuals' subjective intention toward using a technology over a longer period [21]	4.81 (0.34)	72 (97.3)	4.82 (0.39)	125 (98.4)[l]
BI1	Assuming I had access to PedAMINES, I intend to use it.	4.68 (0.60)	69 (93.2)	4.81 (0.41)	126 (99.2)[l]
BI2	I predict I would use PedAMINES in the next 6 months.	4.91 (0.29)	74 (100)	4.85 (0.42)	126 (99.2)[l]
BI3	I expect my use of PedAMINES to continue in the future.	4.84 (0.37)	74 (100)	4.80 (0.44)	125 (98.4)[l]

[a]Presented with the source references from which the items were derived and adapted for the context of this study.

[b]Each item was measured on a 5-point Likert-type scale (1=*strongly disagree* and 5=*strongly agree*); the higher the score, the more agreement with the statement.

[c]PU: perceived usefulness.

[d]PedAMINES: Pediatric Accurate Medication in Emergency Situations.

[e]PEOU: perceived ease of use.

[f]TTF: task–technology fit.

[g]Missing data from nurses, n=1.

[h]PE: performance expectancy.

[i]II: image.

[j]Missing data under nurses, n=2.

[k]N=126.

[l]N=127.

[m]PI: personal innovativeness.

[n]Missing data under nurses, n=1.

[o]A: acceptance.

[p]Missing data under nurses, n=2.

[q]BI: behavioral intention to use.

[r]Missing data under nurses, n=1.

Discussion

Principal Findings

The main finding of our study was that the usability assessment of the PedAMINES app scored high on all items of the SUS questionnaire, with >86% of paramedics and >93% of nurses rating the overall usability of the app as excellent (ie, scoring >80 on the SUS). This suggests that the app is highly usable by users targeted for its purpose and appears to be an accepted supportive tool. This usability observed regardless of years of experience, age, or gender suggests a worthwhile benefit of its use by novice emergency care providers and those with limited exposure to children who are critically ill and few opportunities to prepare emergency medications in pediatric doses, particularly in general hospitals handling pediatric emergencies and in the prehospital setting. Leveling providers' compounding skills could indeed prove to be an advantage in high-stakes clinical events with infrequent occurrences, such as pediatric cardiopulmonary resuscitation. To our knowledge, this is the first study to report usability testing and the intent of paramedics and nurses to use a mobile app as a supportive digital tool for emergency drug preparation in pediatric life-threatening situations. The fact that participants who were previously naive about the PedAMINES app were able to use it correctly to significantly reduce medication error rates [32,33] after a single 5-minute prescenario training and broadly agree on its usability validated the user-centered design [58-60], the Fitts laws [61], and progressive disclosure [62] principles that underpinned its iterative development process [63]. This approach has proven beneficial in allowing end users to influence the development process and increase the final usability of web-based HIT [64,65]. To provide some perspective, the mean total SUS scores >89 in this study were higher than those reported in a comparative study of the top 10 nonmedical mobile apps evaluated by >3500 users [66]. The SUS has proven to be a highly robust and versatile tool for collecting users' subjective evaluations of a product's usability [67]. In high-stakes, critical situations where time is of the essence, the usability of dedicated apps must be high as there is no room at that moment to become familiar with the app [68]. This study suggests that endowing paramedics and nurses with usability-proven mobile apps might contribute to improving the safety of the drug administration process in pediatric emergency care.

To limit and mitigate the likelihood of medication errors, several assistive eHealth technologies have been developed over the past decades to target and support individual medication steps [34]. However, before an eHealth tool such as a mobile app can be adopted and implemented into clinical practice to support the delivery of health care, usability testing and the likelihood that it will be accepted as an aid by its future users are prerequisites for success [69,70]. Usability (ie, the extent to which a system, product, or service can be used by end users to achieve specific goals with effectiveness, efficiency, and satisfaction in a given context of use [71]) is recognized as a key quality factor in determining the success and adoption of an app [26]. Usability testing of an app with end users early in the process may also uncover potential issues related to poor app development and design, which could otherwise ultimately lead to endangering patient safety [72,73]. Although previous research has shown that mHealth apps can help improve the quality and safety of care [74], functional characteristics have often been privileged to the detriment of the needs and characteristics of end users [23]. Consequently, users may be reluctant to adopt them or only use them for a short span of time after their introduction and then abandon them [75,76]. To fully anticipate their acceptance and long-term adoption rate, it is essential to look beyond the technology itself and its usability by also considering end users' beliefs, perceptions, and intentions regarding its use [24].

To date, only a small fraction of mHealth apps have published their usability evaluation results, with most apps developed in the commercial sector that have rarely been published in the scientific literature [26]. In the field of cardiopulmonary resuscitation, a recent study identified 34 available mobile apps on Google Play and Apple App stores [68]. However, many of these apps are marginally medically correct, have not been validated in evidence-based studies, and have limited usability [77]. Of the few available medical mobile apps that offer weight-based drug dosing, such as Handtevy Mobile [78], SafeDose Mobile (eBroselow) [79], Infinite Dose PRO [80], Pedi QuikCalc [81], PEDeDose [82], and EZDrips Peds [83], none have shown proven results of their efficacy, especially in terms of usability.

Although studies have been conducted in recent years to assess user attitudes toward new HIT, there is a lack of research on the perception and acceptance of mHealth technologies by health care providers [84]. In this study, we evaluated 8 technology acceptance constructs to capture the different aspects that could drive the intention to use the PedAMINES app from the perspective of emergency care professionals. We report that paramedics and nurses had an overall acceptant attitude toward the app and agreed with its use as a means of supporting the drug preparation process during emergency care. Among the identified constructs that drive technology acceptance, participants strongly agreed with the usefulness, ease of use, TTF, performance, and behavioral intention to use the app, with mean total scores >4.5/5. In addition, attitudes toward use and acceptance were positive. As these constructs have been shown to be among the most important factors driving individuals to adopt mHealth apps [69,75,85,86], the results of our study provide support for the future adoption of our app by nurses and paramedics in emergency care. However, long-term adoption studies in clinical practice need to be conducted to confirm this. One of the strengths of the study is that the evaluation of the app did not focus on the behavioral intent of its adoption under *traditional* laboratory-based conditions but took into account the usual context of its use by evaluating its usability and acceptability in simulated, in situ conditions very close to the reality in which the app should ultimately be deployed [73].

Our study also found that nurses' attitudes toward app use were slightly negatively influenced by age, whereas usability ratings were not. Consistent with previous research, this may be as some form of resistance may characterize the attitude toward the use of new technologies among nurses with longer experience and their reliance on their own prior learning without

the use of technology to assist them [87-89]. Paramedics were represented in younger age categories than nurses, which may explain the lack of a significant age influence on their attitude toward the app. However, previous studies on the effect of age on attitudes toward novelty have yielded controversial results, and caution should be exercised about this interpretation related to the influence of age, and this aspect requires further study [87]. Of note, our study also found that images appeared to be the least relevant technology acceptance construct. Similar to a previous study [90], this could be explained by the nature of the intervention, as the impact of an app on the image of its users among their peers can presumably only be properly assessed through long-term use in routine practice and not through a single use, as was the case in this study. Future research should determine whether the acceptability of the app, as assessed through this simulation-based study, will translate into its long-term use in real-life situations, as the success of an app likely depends primarily on sustained use over time rather than on its first use [91,92]. Meanwhile, this study generated useful knowledge to guide future developments of other mHealth interventions using in situ high-fidelity simulation and the SUS and technology acceptance self-administered survey questionnaires as a basis for exploring the usability and acceptability of mobile apps at the point of care in emergency medicine.

Limitations

This study has some limitations. First, the app was evaluated in simulation-based trials, and its usability and acceptance in real-life situations might be different. However, the simulated scenarios allowed for consistent assessment of the app among participants in a standardized manner, which would have been difficult to achieve in actual cardiopulmonary resuscitations. In situ clinical simulation, which involves observing representative users performing tasks in their representative environments, has proven to be a valuable way of addressing specific factors that influence technology usability and adoption and ensuring that the results obtained can be generalized to the real world [73]. Second, the survey used to assess the acceptance of the app was not the original UTAUT model but an adapted version with additional valid and reliable constructs. Adding constructs to the original framework is a common practice [93]. As suggested by Venkatesh et al [21], the addition of relevant constructs to the original UTAUT model, where each construct can be interpreted independently of the others, may contribute to extending its applicability to other contexts in an attempt to better understand factors influencing adoption and behavioral intention to use a technology. For example, the TTF construct was used here to evaluate the users' perception of the complexity of the task to be handled by being supported by the app. Third, this study was limited to the evaluation of the app in the context of drug preparations at pediatric doses. It does not provide information on its usability and acceptance if it were to be used to support the preparation of drugs for adults. Fourth, no qualitative usability testing was conducted in the field as part of this study so as to not detract from the scenario and primary outcome. More extensive qualitative laboratory and field testing of the app will be the focus of future research. Finally, the SUS questionnaire measured users' perceived usability of the app rather than its actual usability. Users were not informed of their performance after the scenarios and only filled out the questionnaire based on their perception of their app use. Future studies must consider assessing objective PedAMINES usability metrics, such as task completion rate and efficiency on time on task.

Conclusions

This study evaluated the usability and technology acceptance of a mobile app to assist in the preparation of emergency drugs at the point of care by skilled nurses and advanced paramedics during simulated pediatric cardiopulmonary resuscitations. We found excellent usability and high technology acceptance. This provides information not only about the initial adoption of the app by these caregivers but also, more broadly, about the likelihood of successful adoption of mHealth apps in emergency care that have previously gone through such a usability and technology acceptance evaluation process. Our findings contribute to the exploration of factors influencing the usability and acceptance of mHealth apps by emergency caregivers, particularly when devolved to pediatrics, and lay the groundwork for future clinical practice and long-term research.

Acknowledgments

The authors are grateful to Rosemary Sudan (Freelance Technical Editor) for providing editorial assistance. The authors are also grateful to Mr Lorenz Vogt and Mr Alexandre Glasner for their assistance in translation at some emergency medical service (EMS) centers. They were compensated for their work.

This study was funded by the Swiss National Science Foundation (fund SNSF_32003B_169348 and SNSF_32003B-182374). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

The collaborating members of the PedAMINES (Pediatric Accurate Medication in Emergency Situations) prehospital group are as follows: Philippe Cottet, MD (ACE Geneva Ambulances); Marec Saillant, EMS paramedic (Geneva Team Ambulances); Renaud Grandjean, EMS paramedic (SK Ambulances); Annick Leuenberger, EMS paramedic (Secours Ambulances Genève); Pascal Donnet, EMS paramedic, and Philippe Hauck, EMS paramedic (Service de Sauvetage et de Lutte contre les Incendies Aéroportuaires); Sébastien Pappalardo, EMS paramedic (Service d'Incendie et de Secours); Philippe Nidegger, EMS paramedic (Ambulance Riviera); David Neel, flight EMS paramedic (Air Zermatt SA); Stephan Steinhauser, MD (Höhere Fachschule für Rettungsberufe); Michel Ceschi, EMS paramedic, and Bruno Belli, EMS paramedic (Servizio Ambulanza Locarnese e Valli);

Sébastien Ottet, EMS paramedic, and Wenceslao Garcia, MD (Ambulances du Sud Fribourgeois); Yoan Mollier (EMS paramedic), Yves Vollenweider (EMS paramedic), and Pierre Voumard (EMS paramedic; Service Communal de la Sécurité); Karine Corbat, EMS paramedic, and Philippe Robadey, EMS paramedic (Service de Protection et Sauvetage Lausanne); and Joël Bauer, EMS paramedic, and Cyril Berger, EMS paramedic (Centre de Secours et d'Urgences Morges-Aubonne). These contributors coordinated the trial in their respective EMS centers. None of them received compensation for their role in the study.

The collaborating members of the PedAMINES trial group are as follows: Mario Gehri, MD, and Corine Yersin, registered nurse (RN; HEL, Lausanne University Children's Hospital); Daniel Garcia, MD, and Franziska Hermann Marina, RN (Inselspital, Bern University Children's Hospital); Cosette Pharisa, MD, and Monika Spannaus, RN (Fribourg Children's Hospital); Laurence Racine (MD), Bernard Laubscher (MD), and Carole Vah (RN; Neuchâtel Children's Hospital); Juan Llor, MD, and Aude Juzan, RN (Sion Children's Hospital). These contributors coordinated the trial in their respective pediatric emergency department centers. None of them received compensation for their role in the study.

Authors' Contributions

JNS was responsible for the concept and design of the study, coordinated the overall conduct of the trials at each center, was responsible for the literature search, interpreted the data, prepared the figures and tables, and drafted the manuscript. LG and MS contributed to the drafting of the manuscript and critical review of its content with the support of CC, LS, FR, FE, and SM. LB, SM, and JNS were responsible for the data acquisition. CC performed the data analyses. KH and FH provided technical and material support. FE was responsible for the development of the project software with the support of SM and JNS. CL and AG oversaw the development of the project software. SM was the main trial coordinator. SM procured funding for the prehospital trial, whereas AG procured funding for the in-hospital trial. JNS was responsible for the overall conduct of this study and oversaw the writing of this manuscript. All authors have contributed to, seen, and approved the final submitted version of the manuscript; had full access to all the data, including statistical reports and tables in the study; and take responsibility for the integrity of the data and the accuracy of the data analysis. The corresponding author (JNS) confirms that he had full access to the participants' data and endorsed the final responsibility for the submission. He further affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any deviations from the study plan have been explained.

Conflicts of Interest

Geneva University Hospitals are owners of the PedAMINES (Pediatric Accurate Medication in Emergency Situations) app. The app is currently commercially available on Google Play Store and Apple App Store for research and educational purposes. JNS, CL, AG, FE, and SM declare individual intellectual property rights on this app and, as employees of Geneva University Hospitals, indirect institutional rewarding through its commercialization (ie, without personal enrichment). The authors declare no other relationships or activities that could appear to have influenced the submitted work. All authors have completed the International Committee of Medical Journal Editors uniform disclosure form and declare no support from commercial entities for the submitted work and no financial relationships with any commercial entities that might have an interest in the submitted work in the previous 3 years.

Multimedia Appendix 1
The System Usability Scale questionnaire.
[DOCX File , 17 KB - humanfactors_v9i1e35399_app1.docx]

Multimedia Appendix 2
The technology acceptance survey for the evaluation of health professionals' acceptance of a mobile app (PedAMINES [Pediatric Accurate Medication in Emergency Situations]) for pediatric drug preparation.
[DOCX File , 24 KB - humanfactors_v9i1e35399_app2.docx]

Multimedia Appendix 3
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online TeleHealth) checklist (version 1.6.1).
[PDF File (Adobe PDF File), 415 KB - humanfactors_v9i1e35399_app3.pdf]

Multimedia Appendix 4
Results of the System Usability Scale items for all participants by occupation type (paramedics or nurses).
[DOCX File , 18 KB - humanfactors_v9i1e35399_app4.docx]

Multimedia Appendix 5

Results of the technology acceptance survey items by occupation type (paramedics or nurses) and score ranges.
[DOCX File , 16 KB - humanfactors_v9i1e35399_app5.docx]

References

1. Kruse CS, Beane A. Health information technology continues to show positive effect on medical outcomes: systematic review. J Med Internet Res 2018 Feb 05;20(2):e41 [FREE Full text] [doi: 10.2196/jmir.8793] [Medline: 29402759]
2. Magrabi F, Habli I, Sujan M, Wong D, Thimbleby H, Baker M, et al. Why is it so difficult to govern mobile apps in healthcare? BMJ Health Care Inform 2019 Nov 18;26(1):e100006 [FREE Full text] [doi: 10.1136/bmjhci-2019-100006] [Medline: 31744843]
3. Ryu S. Book review: mHealth: new horizons for health through mobile technologies: based on the findings of the second global survey on ehealth (global observatory for ehealth series, volume 3). Healthc Inform Res 2012;18(3):231. [doi: 10.4258/hir.2012.18.3.231]
4. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
5. Number of apps available in leading app stores as of 1st quarter 2021. Statista. URL: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ [accessed 2022-02-19]
6. Baxter C, Carroll J, Keogh B, Vandelanotte C. Assessment of mobile health apps using built-in smartphone sensors for diagnosis and treatment: systematic survey of apps listed in international curated health app libraries. JMIR Mhealth Uhealth 2020 Feb 03;8(2):e16741 [FREE Full text] [doi: 10.2196/16741] [Medline: 32012102]
7. Jones M, Morris J, Deruyter F. Mobile healthcare and people with disabilities: current state and future needs. Int J Environ Res Public Health 2018 Mar 14;15(3):515 [FREE Full text] [doi: 10.3390/ijerph15030515] [Medline: 29538292]
8. Subhi Y, Bube SH, Rolskov Bojsen S, Skou Thomsen AS, Konge L. Expert involvement and adherence to medical evidence in medical mobile phone apps: a systematic review. JMIR Mhealth Uhealth 2015 Jul 27;3(3):e79 [FREE Full text] [doi: 10.2196/mhealth.4169] [Medline: 26215371]
9. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. Evid Based Med 2013 Jun;18(3):90-92. [doi: 10.1136/eb-2012-100885] [Medline: 22923708]
10. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14 [FREE Full text] [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
11. Labrique A, Vasudevan L, Chang LW, Mehl G. H_pe for mHealth: more "y" or "o" on the horizon? Int J Med Inform 2013 May;82(5):467-469 [FREE Full text] [doi: 10.1016/j.ijmedinf.2012.11.016] [Medline: 23279850]
12. Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]
13. Van Velthoven MH, Smith J, Wells G, Brindley D. Digital health app development standards: a systematic review protocol. BMJ Open 2018 Aug 17;8(8):e022969 [FREE Full text] [doi: 10.1136/bmjopen-2018-022969] [Medline: 30121614]
14. Strudwick G. Predicting nurses' use of healthcare technology using the technology acceptance model: an integrative review. Comput Inform Nurs 2015 May;33(5):189-98; quiz E1. [doi: 10.1097/CIN.0000000000000142] [Medline: 25974361]
15. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
16. King WR, He J. A meta-analysis of the technology acceptance model. Inf Manag 2006 Sep;43(6):740-755. [doi: 10.1016/j.im.2006.05.003]
17. Dishaw MT, Strong DM. Extending the technology acceptance model with task–technology fit constructs. Inf Manag 1999 Jul;36(1):9-21 [FREE Full text] [doi: 10.1016/s0378-7206(98)00101-3]
18. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
19. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decision Sciences 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
20. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
21. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly 2012;36(1):157. [doi: 10.2307/41410412]
22. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
23. 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]
24. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding clinicians' adoption of mobile health tools: a qualitative review of the most used frameworks. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e18072 [FREE Full text] [doi: 10.2196/18072] [Medline: 32442132]

25. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
26. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
27. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. London: Taylor and Francis; Jun 11, 1996.
28. Brooke J. SUS: a retrospective. J Usab Stud 2013 Feb 1;8(2):29-40.
29. Sousa V, Lopez KD. Towards usable e-health. Appl Clin Inform 2017 Dec 21;08(02):470-490. [doi: 10.4338/aci-2016-10-r-0170]
30. Mayer MA, Rodríguez Blanco O, Torrejon A. Use of health apps by nurses for professional purposes: web-based survey study. JMIR Mhealth Uhealth 2019 Nov 1;7(11):e15195. [doi: 10.2196/15195]
31. Kong T, Scott MM, Li Y, Wichelman C. Physician attitudes towards-and adoption of-mobile health. Digit Health 2020;6:2055207620907187 [FREE Full text] [doi: 10.1177/2055207620907187] [Medline: 32128235]
32. Siebert JN, Bloudeau L, Combescure C, Haddad K, Hugon F, Suppan L, Pediatric Accurate Medication in Emergency Situations (PedAMINES) Prehospital Group. Effect of a mobile app on prehospital medication errors during simulated pediatric resuscitation: a randomized clinical trial. JAMA Netw Open 2021 Aug 02;4(8):e2123007 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.23007] [Medline: 34459905]
33. Siebert JN, Ehrler F, Combescure C, Lovis C, Haddad K, Hugon F, PedAMINES Trial Group. A mobile device application to reduce medication errors and time to drug delivery during simulated paediatric cardiopulmonary resuscitation: a multicentre, randomised, controlled, crossover trial. Lancet Child Adolesc Health 2019 May;3(5):303-311. [doi: 10.1016/S2352-4642(19)30003-3] [Medline: 30797722]
34. Ehrler F, Siebert JN. PedAMINES: a disruptive mHealth app to tackle paediatric medication errors. Swiss Med Wkly 2020 Aug 24;150:w20335 [FREE Full text] [doi: 10.4414/smw.2020.20335] [Medline: 32920794]
35. Lacour M, Bloudeau L, Combescure C, Haddad K, Hugon F, Suppan L, PedAMINES Prehospital Group. Impact of a mobile app on paramedics' perceived and physiologic stress response during simulated prehospital pediatric cardiopulmonary resuscitation: study nested within a multicenter randomized controlled trial. JMIR Mhealth Uhealth 2021 Oct 07;9(10):e31748 [FREE Full text] [doi: 10.2196/31748] [Medline: 34617916]
36. Siebert JN, Bloudeau L, Ehrler F, Combescure C, Haddad K, Hugon F, et al. A mobile device app to reduce prehospital medication errors and time to drug preparation and delivery by emergency medical services during simulated pediatric cardiopulmonary resuscitation: study protocol of a multicenter, prospective, randomized controlled trial. Trials 2019 Nov 20;20(1):634 [FREE Full text] [doi: 10.1186/s13063-019-3726-4] [Medline: 31747951]
37. Siebert JN, Ehrler F, Lovis C, Combescure C, Haddad K, Gervaix A, et al. A mobile device app to reduce medication errors and time to drug delivery during pediatric cardiopulmonary resuscitation: study protocol of a multicenter randomized controlled crossover trial. JMIR Res Protoc 2017 Aug 22;6(8):e167 [FREE Full text] [doi: 10.2196/resprot.7901] [Medline: 28830854]
38. Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, International Network for Simulation-based Pediatric Innovation, Research,Education (INSPIRE) Reporting Guidelines Investigators. Reporting guidelines for health care simulation research: extensions to the CONSORT and STROBE statements. Simul Healthc 2016 Aug;11(4):238-248. [doi: 10.1097/SIH.0000000000000150] [Medline: 27465839]
39. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
40. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010 Mar 23;340:c332 [FREE Full text] [doi: 10.1136/bmj.c332] [Medline: 20332509]
41. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. J Med Internet Res 2020 Jul 06;22(7):e17256 [FREE Full text] [doi: 10.2196/17256] [Medline: 32628122]
42. Gao M, Kortum P, Oswald FL. Multi-language toolkit for the system usability scale. Int J Human Comput Interaction 2020 Aug 19;36(20):1883-1901. [doi: 10.1080/10447318.2020.1801173]
43. Gronier G, Baudet A. Psychometric evaluation of the F-SUS: creation and validation of the French version of the system usability scale. Int J Human Comput Interaction 2021 Mar 17;37(16):1571-1582. [doi: 10.1080/10447318.2021.1898828]
44. Acceptance factors and use of information technologies: an empirical study on the use of "Rational Suite" software by the employees of a large IT services company. Quebec University (UQAM). URL: https://archipel.uqam.ca/3960/1/M11905.pdf [accessed 2022-02-20]
45. Heinsch M, Wyllie J, Carlson J, Wells H, Tickner C, Kay-Lambkin F. Theories informing eHealth implementation: systematic review and typology classification. J Med Internet Res 2021 May 31;23(5):e18500 [FREE Full text] [doi: 10.2196/18500] [Medline: 34057427]
46. Goodhue DL, Thompson RL. Task-technology fit and individual performance. MIS Quarterly 1995 Jun;19(2):213. [doi: 10.2307/249689]

47. Strong DM, Dishaw M, Bandy D. Extending task technology fit with computer self-efficacy. In: Eighth Americas Conference on Information Systems. 2002 Presented at: Eighth Americas Conference on Information Systems; Aug 8-11, 2002; Dallas, United States. [doi: 10.1145/1161345.1161358]
48. Klopping IM, McKinney E. Extending the technology acceptance model and the task-technology fit model to consumer e-commerce. Inform Technol Learn Perform J 2004;22(1):35-48.
49. Moore GC, Benbasat I. Development of an instrument to measure the perceptions of adopting an information technology innovation. Inf Syst Res 1991 Sep;2(3):192-222. [doi: 10.1287/isre.2.3.192]
50. Yi M, Jackson J, Park J, Probst J. Understanding information technology acceptance by individual professionals: toward an integrative view. Inf Manag 2006 Apr;43(3):350-363 [FREE Full text] [doi: 10.1016/j.im.2005.08.006]
51. 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]
52. ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. Stat Med 1999 Aug 15;18(15):1905-1942. [Medline: 10532877]
53. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 May 01;4(3):114-123. [doi: 10.5555/2835587.2835589]
54. Sauro J. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Scotts Valley, California, US: CreateSpace Independent Publishing; 2011.
55. Dwivedi YK, Rana NP, Jeyaraj A, Clement M, Williams MD. Re-examining the unified theory of acceptance and use of technology (UTAUT): towards a revised theoretical model. Inf Syst Front 2017 Jun 8;21(3):719-734. [doi: 10.1007/s10796-017-9774-y]
56. Klopping I, McKinney E. Extending the technology acceptance model and the task-technology fit model to consumer e-commerce. Inform Technol Learn Perform J 2004;22(1):35-48.
57. Agarwal R, Prasad J. A conceptual and operational definition of personal innovativeness in the domain of information technology. Inf Syst Res 1998 Jun;9(2):204-215. [doi: 10.1287/isre.9.2.204]
58. Durrani S, Durrani QS. Applying cognitive psychology to user interfaces. In: Proceedings of the First International Conference on Intelligent Human Computer Interaction. 2009 Presented at: Proceedings of the First International Conference on Intelligent Human Computer Interaction; Jan 20–23, 2009; Allahabad, India URL: https://doi.org/10.1007/978-81-8489-203-1_14 [doi: 10.1007/978-81-8489-203-1_14]
59. Oviatt S. Human-centered design meets cognitive load theory: designing interfaces that help people think. In: Proceedings of the 14th ACM international conference on Multimedia. 2006 Presented at: Proceedings of the 14th Association for Computing Machinery (ACM) International Conference on Multimedia; Oct 23-27, 2006; Santa Barbara, United States. [doi: 10.1145/1180639.1180831]
60. Harper S, Michailidou E, Stevens R. Toward a definition of visual complexity as an implicit measure of cognitive load. ACM Trans Appl Percept 2009 Feb 15;6(2):1-18. [doi: 10.1145/1498700.1498704]
61. MacKenzie IS. Fitts' law as a research and design tool in human-computer interaction. Human Comput Interact 2009 Nov 11;7(1):91-139. [doi: 10.1207/s15327051hci0701_3]
62. Spillers F. Progressive disclosure. In: The Glossary of Human Computer Interaction. Copenhagen: Interaction Design Foundation; 2016.
63. Hagberg H, Siebert J, Gervaix A, Daehne P, Lovis C, Manzano S, et al. Improving drugs administration safety in pediatric resuscitation using mobile technology. Stud Health Technol Inform 2016;225:656-657. [Medline: 27332296]
64. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
65. Saparamadu AA, Fernando P, Zeng P, Teo H, Goh A, Lee JM, et al. User-centered design process of an mHealth app for health professionals: case study. JMIR Mhealth Uhealth 2021 Mar 26;9(3):e18079 [FREE Full text] [doi: 10.2196/18079] [Medline: 33769297]
66. Kortum P, Sorber M. Measuring the usability of mobile applications for phones and tablets. Int J Human Comput Interaction 2015 Jul 31;31(8):518-529. [doi: 10.1080/10447318.2015.1064658]
67. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Human Comput Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
68. Metelmann B, Metelmann C, Schuffert L, Hahnenkamp K, Brinkrolf P. Medical correctness and user friendliness of available apps for cardiopulmonary resuscitation: systematic search combined with guideline adherence and usability evaluation. JMIR Mhealth Uhealth 2018 Nov 06;6(11):e190 [FREE Full text] [doi: 10.2196/mhealth.9651] [Medline: 30401673]
69. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220 [FREE Full text] [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
70. Philippi P, Baumeister H, Apolinário-Hagen J, Ebert DD, Hennemann S, Kott L, et al. Acceptance towards digital health interventions - Model validation and further development of the Unified Theory of Acceptance and Use of Technology. Internet Interv 2021 Dec;26:100459 [FREE Full text] [doi: 10.1016/j.invent.2021.100459] [Medline: 34603973]

71. Bevan N. International standards for HCI and usability. Int J Human Comput Stud 2001 Oct;55(4):533-552 [FREE Full text] [doi: 10.1006/ijhc.2001.0483]
72. Dexheimer JW, Kurowski BG, Anders SH, McClanahan N, Wade SL, Babcock L. Usability evaluation of the SMART application for youth with mTBI. Int J Med Inform 2017 Jan;97:163-170 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.10.007] [Medline: 27919376]
73. Kushniruk A, Nohr C, Jensen S, Borycki EM. From usability testing to clinical simulations: bringing context into the design and evaluation of usable and safe health information technologies. Yearb Med Inform 2018 Mar 05;22(01):78-85. [doi: 10.1055/s-0038-1638836]
74. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020 Jan 13;3(1):4 [FREE Full text] [doi: 10.1038/s41746-019-0206-x] [Medline: 31970289]
75. Binyamin SS, Zafar BA. Proposing a mobile apps acceptance model for users in the health area: a systematic literature review and meta-analysis. Health Informatics J 2021;27(1):1460458220976737 [FREE Full text] [doi: 10.1177/1460458220976737] [Medline: 33438494]
76. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
77. Lauridsen KG, Nadkarni VM, Berg RA. Man and machine: can apps resuscitate medical performance? Lancet Child Adolescent Health 2019 May;3(5):282-283. [doi: 10.1016/s2352-4642(19)30032-x]
78. Safer dosing stronger documentation. Handtevy. URL: https://www.handtevy.com [accessed 2022-02-20]
79. Are you sure you're doing the right thing for your patients when it comes to medication delivery? SafeDose. URL: https://www.safedoseinc.com [accessed 2022-02-20]
80. Infinite Dose PRO: the smart dosage calculator. Google Play. URL: https://play.google.com/store/apps/details?id=com.moaness.InfiniteDose.pro&hl=iw&gl=US [accessed 2022-02-20]
81. Welcome to Pedi QuikCalc!. Pedi QuikCalc. URL: http://www.pediquikcalc.com/PediQuikCalc/Welcome.html [accessed 2022-02-20]
82. Higi L, Käser K, Wälti M, Grotzer M, Vonbach P. Description of a clinical decision support tool with integrated dose calculator for paediatrics. Eur J Pediatr 2022 Feb 15;181(2):679-689. [doi: 10.1007/s00431-021-04261-2] [Medline: 34524516]
83. EZDrips homepage. EZDrips. URL: https://www.ezdrips.com/fr/ [accessed 2022-02-20]
84. Sezgin E, Özkan-Yildirim S, Yildirim S. Understanding the perception towards using mHealth applications in practice. Inf Development 2016 Dec 19;34(2):182-200. [doi: 10.1177/0266666916684180]
85. Woldeyohannes HO, Ngwenyama OK. Factors influencing acceptance and continued use of mHealth apps. In: HCI in Business, Government and Organizations. Cham: Springer International Publishing; 2017.
86. Wang C, Qi H. Influencing factors of acceptance and use behavior of mobile health application users: systematic review. Healthcare (Basel) 2021 Mar 22;9(3):357 [FREE Full text] [doi: 10.3390/healthcare9030357] [Medline: 33809828]
87. Assaker G. Age and gender differences in online travel reviews and user-generated-content (UGC) adoption: extending the technology acceptance model (TAM) with credibility theory. J Hospit Market Manag 2019 Aug 28;29(4):428-449. [doi: 10.1080/19368623.2019.1653807]
88. Song L, Park B, Oh K. Analysis of the technology acceptance model in examining hospital nurses' behavioral intentions toward the use of bar code medication administration. Comput Inform Nurs 2015 Apr;33(4):157-165. [doi: 10.1097/CIN.0000000000000143] [Medline: 25816212]
89. Timmons S. Nurses resisting information technology. Nurs Inq 2003 Dec;10(4):257-269. [doi: 10.1046/j.1440-1800.2003.00177.x] [Medline: 14622372]
90. Magsamen-Conrad K, Dillon JM. Mobile technology adoption across the lifespan: a mixed methods investigation to clarify adoption stages, and the influence of diffusion attributes. Comput Human Behav 2020 Nov;112:106456 [FREE Full text] [doi: 10.1016/j.chb.2020.106456] [Medline: 32834465]
91. Hsiao J, Chen R. Understanding determinants of health care professionals' perspectives on mobile health continuance and performance. JMIR Med Inform 2019 Mar 18;7(1):e12350 [FREE Full text] [doi: 10.2196/12350] [Medline: 30882353]
92. Akter S, Ray P, D'Ambra J. Continuance of mHealth services at the bottom of the pyramid: the roles of service quality and trust. Electron Markets 2012 May 8;23(1):29-47. [doi: 10.1007/s12525-012-0091-5]
93. 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]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online TeleHealth
EMS: emergency medical service
HIT: health information technology

mHealth: mobile health
OHCA: out-of-hospital cardiac arrest
PedAMINES: Pediatric Accurate Medication in Emergency Situations
REDCap: Research Electronic Data Capture
RN: registered nurse
SUS: System Usability Scale
TAM: Technology Acceptance Model
TTF: task–technology fit
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 03.12.21; peer-reviewed by BJ Nievas-Soriano, M Auerbach; comments to author 24.12.21; revised version received 03.01.22; accepted 20.01.22; published 01.03.22.

Please cite as:
Siebert JN, Gosetto L, Sauvage M, Bloudeau L, Suppan L, Rodieux F, Haddad K, Hugon F, Gervaix A, Lovis C, Combescure C, Manzano S, Ehrler F, PedAMINES Trial Group, PedAMINES Prehospital Group
Usability Testing and Technology Acceptance of an mHealth App at the Point of Care During Simulated Pediatric In- and Out-of-Hospital Cardiopulmonary Resuscitations: Study Nested Within 2 Multicenter Randomized Controlled Trials
JMIR Hum Factors 2022;9(1):e35399
URL: https://humanfactors.jmir.org/2022/1/e35399
doi:10.2196/35399
PMID:35230243

Original Paper

REDCap Delivery of a Web-Based Intervention for Patients With Voice Disorders: Usability Study

Danielle Mollie Stambler[1*], MLiS; Erin Feddema[2*], MPH; Olivia Riggins[1*], BS; Kari Campeau[3*], PhD; Lee-Ann Kastman Breuch[1*], PhD; Molly M Kessler[1*], PhD; Stephanie Misono[2*], MPH, MD

[1]Department of Writing Studies, University of Minnesota, Minneapolis, MN, United States
[2]Department of Otolaryngology, University of Minnesota, Minneapolis, MN, United States
[3]Department of English, University of Colorado-Denver, Denver, CO, United States
[*]all authors contributed equally

Corresponding Author:
Danielle Mollie Stambler, MLiS
Department of Writing Studies
University of Minnesota
Nolte Center
315 Pillsbury Dr SE
Minneapolis, MN, 55455
United States
Phone: 1 612 624 3445
Email: stambler@umn.edu

Abstract

Background: Web-based health interventions are increasingly common and are promising for patients with voice disorders because web-based participation does not require voice use. To address needs such as Health Insurance Portability and Accountability Act compliance, unique user access, the ability to send automated reminders, and a limited development budget, we used the Research Electronic Data Capture (REDCap) data management platform to deliver a patient-facing psychological intervention designed for patients with voice disorders. This was a novel use of REDCap.

Objective: We aimed to evaluate the usability of the intervention, with this intervention serving as a use case for REDCap-based patient-facing interventions.

Methods: We used REDCap survey instruments to develop the web-based voice intervention modules, then conducted usability evaluations using (1) heuristic evaluations by 2 evaluators, and (2) formal usability testing with 7 participants, consisting of predetermined tasks, a think-aloud protocol, ease-of-use measurements, a product reaction card, and a debriefing interview.

Results: Heuristic evaluations found strengths in visibility of system status and real-world match, and weaknesses in user control and help documentation. Based on this feedback, changes to the intervention were made before usability testing. Overall, usability testing participants found the intervention useful and easy to use, although testing revealed some concerns with design, content, and terminology. Some concerns were readily addressed, and others required adaptations within REDCap.

Conclusions: The REDCap version of a complex web-based patient-facing intervention performed well in heuristic evaluation and formal usability testing. REDCap can effectively be used for patient-facing intervention delivery, particularly if the limitations of the platform are anticipated and mitigated.

(JMIR Hum Factors 2022;9(1):e26461) doi:10.2196/26461

KEYWORDS

web-based intervention; REDCap; voice disorders; usability study; heuristics; eHealth; online; health; web-based participation; patients; web-based platform

Introduction

Patients and providers are increasingly turning to digital platforms for medical information and support. The COVID-19 pandemic has further reinforced the need for information and intervention delivery that does not require in-person contact. Web-based interventions are particularly effective for disorders that may impact interpersonal interactions because web-based

interventions can reduce barriers to access and communication. One such disorder—vocal dysfunction (or dysphonia)—is common [1] and leads to approximately US $2 billion in annual loss in work productivity [2-4], as well as significant functional and social impairments. Lower quality of life (voice-related) has been reported by patients with lower perceived control [5] (ie, perceived control over events or one's reactions to events). Perceived control can be increased through targeted web-based intervention [6,7], and greater perceived control is associated with better patient-reported, disease-specific, and overall outcomes such as depression, diabetes, asthma, heart disease, and blood pressure [8-13]. A web-based psychological intervention could thus enhance voice treatment outcomes in a low-cost, accessible way.

Web-based interventions are an especially promising avenue for patients with voice disorders because web-based participation does not require voice use. To be usable by patients, a web-based platform should be Health Insurance Portability and Accountability Act (HIPAA)–compliant, with unique access for users, the ability to send automated reminders, and effective usability. We initially developed a custom website for this purpose, and initial intervention findings were promising [14], but we experienced difficulties related to cost, transparency, and troubleshooting timeliness. Those difficulties were perhaps inevitable, given the limited project resources, and prompted us to explore other potential HIPAA-compliant options for future interventions.

Because our team frequently uses REDCap for data collection, we considered its usefulness for the delivery of a patient-facing web-based intervention. REDCap is an electronic data capture platform widely used in biomedical research because it is secure, HIPAA–compliant, facilitates data exports for analysis easily, and is free or low cost for university researchers under institutional contracts [15]. REDCap's user-friendly interface reduces the need for programming knowledge and provides additional control through customization options. The platform is well supported, with continuous monitoring, systematic updates, and an increasing list of capabilities that provide a near maintenance-free infrastructure to researchers. Our intervention delivered educational modules over a period of time by sending automated email reminders, for which REDCap's survey functionality and Automatic Survey Invitation tool seemed well suited. Enabling modules, with respect to timing and sequence, is dependent on multiple inputs, which REDCap is able to capture, calculate, and modify throughout a participant's use. REDCap functionality also facilitated parallel designs for multiple study arms and simplified study management. We reasoned that the existing functionality might, therefore, be effectively adapted to deliver a patient-facing intervention.

The use of REDCap as a patient-facing intervention is relatively novel. Although REDCap is used by thousands of teams to collect data, it is used far less frequently in a patient-facing manner. The literature does include a few patient-facing studies [16-19], in which REDCap has been used for data collection to assess other (non-REDCap) custom apps or websites. It has also been used for patient interventions as a back-end system paired with custom interfaces such as web pages or interactive forms presented on investigators' tablet computers [20-23]. However, we found only one study [24] on the use of REDCap for patient-facing intervention delivery and usability.

Given the proliferation of study teams utilizing REDCap and concurrent increasing interest in web-based interventions, we aimed to rigorously evaluate the usability of REDCap for patient-facing intervention delivery in a use case. The objectives of this study were to evaluate the usability of the voice intervention within REDCAP using (1) heuristic evaluation and (2) formal usability testing, which were chosen because they generate complementary forms of usability data. In fact, the combined approach—heuristic evaluation and usability testing—has been described as a "1+2 punch" [25]—providing distinct yet complementary data that can form an excellent baseline for usability. Herein, we also suggest strategies for adapting REDCap to patient-facing health interventions.

Methods

REDCap Intervention

The intervention consisted of 3 parts. The first part delivered baseline assessment measures followed by an educational module with instructional videos and self-led exercises. In the second part, participants were invited to complete check-in modules twice a week for up to 3 weeks (Figures 1-3). The third part delivered end-of-study assessment measures (the same as baseline measures) followed by a participation feedback section. The website could adjust total participation time, allowing enough time for the baseline educational module, 2 check-ins, and the final survey module prior to participants starting voice therapy. Individual survey instruments were developed within the REDCap project for each intervention module, and a database instrument was used to set up participant profiles. Conditional logic, using dates and indicator variables manually entered in or captured throughout the intervention, was used to trigger ASIs to alert participants to available modules.

Developing the intervention in REDCap was an iterative process, because REDCap routinely updates functionalities (ie, fixing issues and making desired features possible or easier to implement). The intervention was developed in REDCap (Table 1) with the knowledge that we would later add a comparison arm to be used in a randomized controlled study of the intervention, making use of the randomization tool. The *longitudinal project* setting was tested, but ultimately not used, in this version of the intervention due to limitations of its use with the randomization tool and survey piping features.

Figure 1. Welcome page for the Voice Education Program intervention.

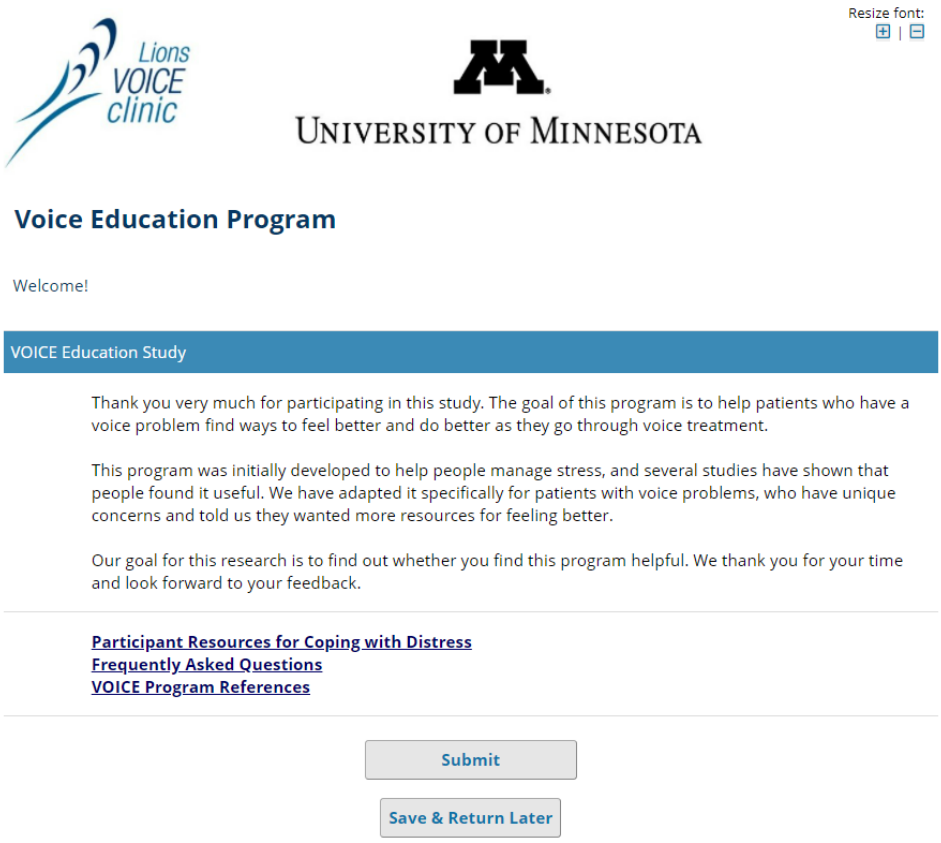


Figure 2. Check-in questionnaire page.

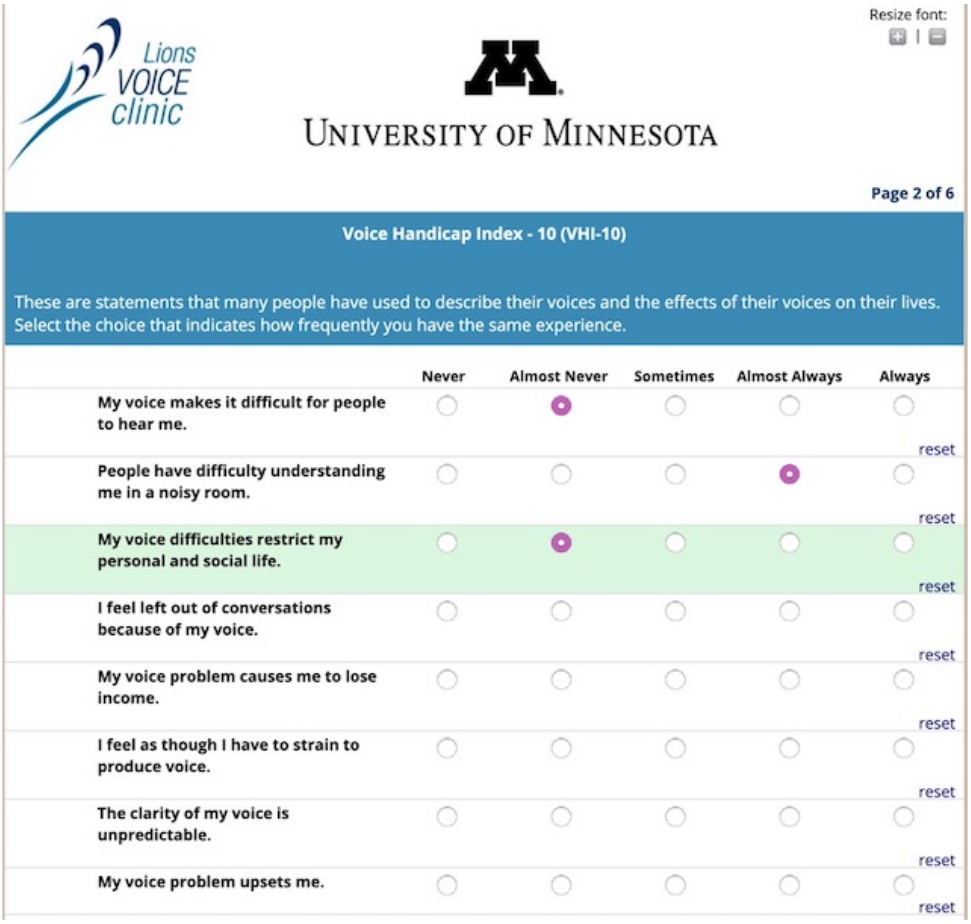


Figure 3. Voice tips page including embedded YouTube video.

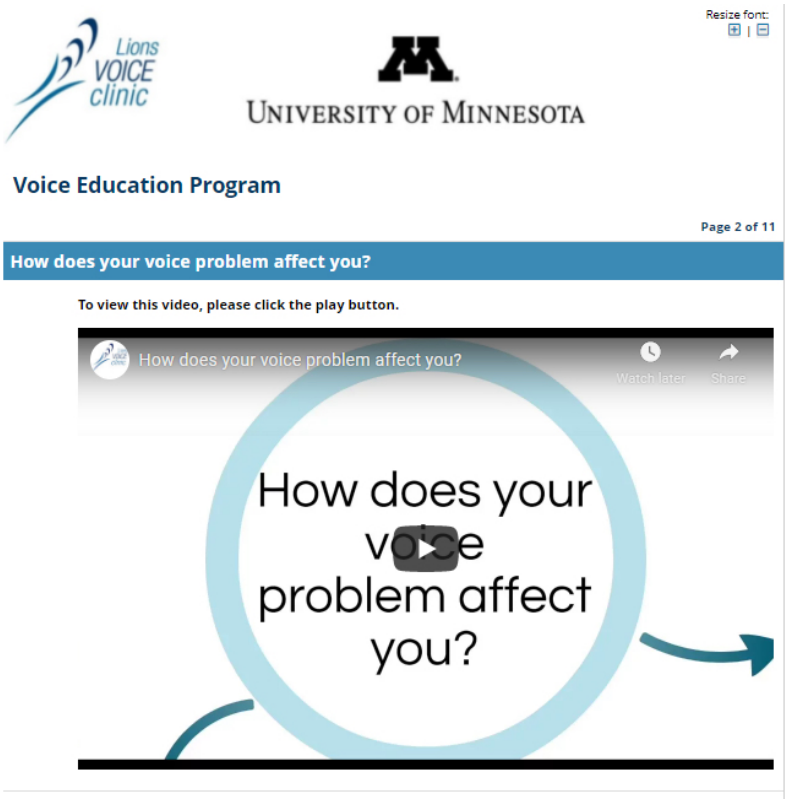


Table 1. Design parameters used in developing the REDCap intervention.

Category and required design parameter	REDCap features[a]
Participant access to intervention	
Intervention accessible through links sent by email. Reminder emails sent with the deadlines for finishing the module.	REDCap generated unique URLs for each participant when sent by *Automatic Survey Invitation* using *Smart Variables*. Dates were *piped* into the email text.
Participants have unique logins to access their information. Participants can save, leave, and return to the website.	Enabled Save & Return in the *Survey Settings*
Data must be kept in a secure, HIPAA[b] compliant database.	REDCap provided a secure web interface and included multiple features to support HIPAA compliance. Both the website and database were housed on secure servers maintained by the researcher's institution.
Intervention design	
Educational material able to be delivered by text and videos.	Used *descriptive text fields*, including in-line video
Self-led exercises present participants' prior responses for reflection and goal-setting.	Previous responses were *piped* into descriptive text for participant review.
The length of intervention can be shortened if the participant's therapy start date is less than 3 weeks away. Module start and completion dates are otherwise used to enable future modules and trigger emails to send.	Therapy date was entered in the participant set up instrument. Start and completion dates were captured as validated date text variables using the *@HIDDEN-SURVEY Action Tag*. Conditional logic in *Automatic Survey Invitations* using *datediff* calculations and indicator variables sent emails. The *Survey Queue* was used when there was not a lag in time between instruments.
Intervention modules are disabled after a period of time to ensure that they are completed in order.	*Time Limit for Survey Completion* option was used in the *Survey Settings*.
Hyperlinks to voice and psychological health tips are embedded in the intervention check-ins. Menu and navigation have hyperlinks to resources including the study FAQ, program references, and supplemental mental health resources.	Hardcoded hyperlink embedded in a survey field connected to another REDCap project using the project's public link. Because a left-hand menu bar was not possible in REDCap and layout was limited to one center panel of text, the menu links were listed in a descriptive text variable at the bottom of each survey page.
Study execution	
The intervention can adapt to changing participant inputs throughout participant, such as:	A project instrument was used for participant setup to enter dates used for *Automatic Survey Invitations* and manage participant information *piped* into the intervention.
Updates in contact information.	
Changing dates of scheduled medical treatment, which affect study participation duration.	An email variable was set as the survey-specific email invitation field in the *Survey Settings* instead of using the *Survey Distribution Tools* to allow for updates over time and ease in participant set up.
	Automatic Survey Invitations were triggered to send the day before the enabling date to allow for date changes during the study, but the researcher could update and retrigger emails if needed using the *Survey Invitation Log*.
The intervention moves through modules in sequence.	Default use of the *Survey Queue* allowed linear progression through surveys.
Participants cannot go backwards and change answers, which is important for data integrity.	Instrument *Survey Settings* were set to prevent return to and modification of completed responses.
Website collects date and time stamps for all responses.	REDCap captured and could report timestamps for survey start, completion, and all responses.

[a]REDCap-specific terms are in italics.

[b]HIPAA: Health Insurance Portability and Accountability Act.

Heuristic Evaluation

A heuristic evaluation is a type of inspection method in which web-based interfaces are evaluated based on a list of guidelines for effective interface design. Our team used well-known heuristics [26] that have been used to evaluate both software and web-based interfaces. The goal of heuristic evaluation is to identify areas in which a web-based intervention meets or does not meet widely accepted guidelines for interface design.

Two usability research assistants who were not involved in intervention development conducted the heuristic evaluation. Working independently, the evaluators identified strengths and weaknesses in the intervention's usability by completing tasks as an end user might, using Nielsen's usability heuristics for interface design [26].

The evaluators then each generated a report that included specific examples of the intervention's degree of compliance

with each usability heuristic, screenshots of the intervention modules, usability rankings, and rationales for rankings. Strengths and weaknesses identified in both reports were used to update the intervention and guided the development of the usability testing plan and testing scenarios. Each evaluator's rankings for the intervention's compliance with the 10 heuristics were standardized on a scale from 1 to 4, with 1 representing ineffective and 4 representing highly effective.

Usability Testing

In addition to the heuristic evaluation, a usability test was conducted to study the interaction of representative users with the web-based intervention. We designed a usability test around key tasks in the web-based intervention and gathered quantitative and qualitative data to identify successes and problem areas as well as overall participant impressions.

Setting

Usability testing was conducted on campus. We began usability testing in the Usability Lab on campus, where high-quality data could be collected, including audiovisual recording and screen captures, and where 2 rooms and a one-way mirror provided an optimal research environment for observers. Complications arose in terms of site location—the campus was unfamiliar to most participants and campus sporting events caused disruptions to driving routes and parking availability—therefore, for the last 2 tests, we moved to a research suite closer to the voice clinic familiar to participants, although the suite did not include high-quality data recording or a one-way mirror. We have discussed these challenges in greater detail elsewhere [27].

Participants and Recruitment

Participants were recruited from MHealth Fairview otolaryngology clinics. Inclusion criteria were diagnosis of a voice disorder; Voice Handicap Index–10 score greater than or equal to 11 [28]; age 18-80 years old; English literacy; and ability to complete informed consent. Because usability testing was completed on campus, potential participants who lived close to the testing site were preferentially invited, although residence location was not used as a strict screening criterion. We recruited 10 participants based on these criteria, and 7 participants completed the usability test.

Ethics

The study was approved by the University of Minnesota institutional review board (1507S75003).

Measures, Procedures, and Analysis

Our study's research questions asked how well patients were able to navigate into, throughout, and exit the intervention; how well patients were able to use multiple choice features to answer questions; and how patients used the *help and documentation* in the FAQ. The usability test consisted of (1) logging into the intervention within REDCap, (2) completing part of a module, (3) reviewing supplementary material and navigating back to a module, (4) exiting REDCap and logging back in, and (5) identifying the help page within the intervention (Multimedia Appendix 1). A moderator facilitated each usability test and asked participants to complete each task using a think-aloud protocol (in which participants share their thoughts as they work through the task [25]). An additional research question, which focused on understanding patient experiences with the intervention, was added during usability testing [27].

Usability measures were time-on-task, task completion rates, issue rates and severities, and subjective user satisfaction. Time-on-task reflected how long it took a participant to complete a task from the time it was given until the time the participant indicated completion. Our goal with time-on-task measures was to establish a realistic baseline time, thus we did not set specific target times. Task completion rates were measured as the percentage of test participants who were able to successfully complete the task without requiring assistance or encountering high-severity issues. Our goal for task completion was 100%. After each usability test, we counted issues and rated each for severity; a high-severity issue (a severity rating of 1) prevented a participant from correctly completing a task, while a low-severity issue (a severity rating of 5) did not change the outcome of the task but resulted in the task being completed less efficiently. Our goals for issues per participant were less than 1 high-severity issue, less than 5 moderate-severity issues, and less than 5 low-severity issues. Additionally, we ranked issues based on their frequency across participants, with low-impact but high-frequency issues overall being rated at a higher severity level.

User satisfaction was measured by asking participants to rate ease of use for each task on a scale of 1 to 5, with 1 representing very easy and 5 representing very difficult. Our goals were to have no posttask user satisfaction rating higher than 3 for any individual participant, and an overall average rating of 2 (easy) for each task across all participants. After each task, the moderator invited participants to offer any comments about the rating they chose.

After each participant completed all tasks, we asked them to complete a product reaction card and debriefing interview describing their experience. The product reaction card was a sheet with a set of 63 positive and negative words from which participants were asked to choose 5 that best described their experience [25]. This set was derived from a desirability matrix [29] of 118 words, with a ratio of 60% positive to 40% negative or neutral words [25]. The matrix can be used in full or abbreviated to gather quick descriptive feedback about participant impressions [30]. Our goal was to have at least 60% of all reaction words be positive. Debriefing interviews included 5 open-ended questions asking participants to describe their initial impressions, how those impressions changed as they used the intervention, what they liked least and best, and what they would change if they could (Multimedia Appendix 2). In combination with the think-aloud protocol, the debriefing interview allowed for insights into participants' health care-related contexts of use and engagement with intervention content [27].

Results

Heuristic Evaluation

The heuristic evaluation (Table 2) indicated that 3 categories could be improved: user control and freedom, consistency and

standards, and help and documentation. These 3 areas of improvement were used to make initial revisions to the intervention and also informed the creation of usability testing tasks and questions.

Table 2. Heuristic evaluation results.

Heuristic	Rating	Strengths	Areas for improvement
1. Visibility of system status	3	• System showed current status effectively in main survey sections via page counts, color confirmation, and written confirmation	• System status and future options were not as apparent in additional help sections
2. Real-world match	2	• Survey section numbering, sequence, and naming were logical and consistent with real world conventions • Survey questions follow conventions for type and format • Embedded YouTube videos take advantage of familiar features, platform	• Procedure for leaving and returning was not conventional or natural • "Survey" terms in standardized research questionnaires did not match real-world conventions • Additional section links are hard to find, and function in unconventional and nonnatural ways
3. User control	2	• Reset function was an effective undo feature	• "Emergency exits" were unclear in additional help sections
4. Consistency	2	• Main survey sections used consistent layout, functioning, color, and terms	• Pop-up boxes, formatting, headers, and tone were inconsistent within and across pages
5. Error prevention	3	• Several effective error prevention features (eg, prohibits leaving questions unanswered)	• No prevention against accidentally closing whole survey window without saving
6. Recognition	3	• Instructions for system use are readily available throughout • Questionnaires and check-ins provide built-in references to past information • Educational videos, FAQ[a], and additional resources are available at the bottom of each page	• Contents and options are not centrally listed in the additional vocal health tips sections
7. Flexibility	3	• Font resize option is available • Survey queue automatically accordions up as surveys are completed, but still provides an option to view all	• Menus with links to additional resources and vocal health tips sections cannot be hidden
8. Aesthetic	3	• Aesthetic is simple, neutral, and uncluttered	• Some images in the additional vocal health tips sections are less relevant and therefore less impactful than they could be
9. Error messaging	3	• Error messaging is clear and provides both an explanation and a solution	• None identified
10. Help and documentation	2	• FAQ and help email are readily available on survey queue/home page • FAQ is available on all main survey pages • Instructions for use are available throughout the module	• Help is not searchable • No documentation for technical issues • No centralized overview of instructions, features, problems, and complex tasks

[a]FAQ: frequently asked question.

Heuristic evaluation indicated that visibility of system status, real-world match, and recognition were all intervention strengths, aligning with REDCap's ability to provide a stable platform that matches users' expectations for websites, without requiring any specialized development knowledge to build and maintain.

The heuristic evaluation indicated user control and freedom, consistency and standards, and help and documentation as 3 major heuristic categories in need of improvement. Of these weaknesses, the category user control and freedom was most pertinent to REDCap's functionality as a platform. Consistency and standards, as well as content in help and documentation, were readily addressable once identified. This included following recommendations for plain language [31] and ensuring parallel structure. REDCap allowed for immediate updating of all edited content without the need to rely on a third party for content editing.

Usability Testing

A total of 10 participants were recruited; 7 participants completed the usability test. All participants were patients at MHealth/Fairview with voice problems (Table 3).

Table 3. Demographic characteristics of study participants.

Characteristic	Value
Age (years), mean (range)	51 (30-71)
Gender (n=7), n (%)	
Male	1 (14)
Female	5 (71)
Gender nonconforming	1 (14)
Race/ethnicity (n=7), n (%)	
White	5 (71)
African American	1 (14)
Asian American	1 (14)
Education (n=7), n (%)	
Some college credits, no degree	1 (14)
Bachelor's degree	1 (14)
Graduate degree	5 (71)
Used web-based health resources before (n=7), n (%)	
At least once a week	2 (28)
At least once a month	1 (14)
Less than once a month	3 (43)
Decline to answer	1 (14)

Results from the usability tests include time-on-task, task completion rates, issue rates and severity, product reaction card selection, and qualitative data from debriefing interviews. Task completion rates ranged from 71% (5/7) to 100% (7/7). Overall, mean posttask ratings were close to our goals, and the individual highest posttask ratings exceeded goals on 4 of 5 tasks (Figure 4).

Issue rates met specified goals for all but 1 participant, who experienced several critical issues; only 1 issue—the use of the *Submit* button—reflected an issue both high in frequency and severity. Four other issues were noted as high impact but with low severity (Table 4).

Figure 4. Time-on-task, mean post-task ratings, and task completion rates.

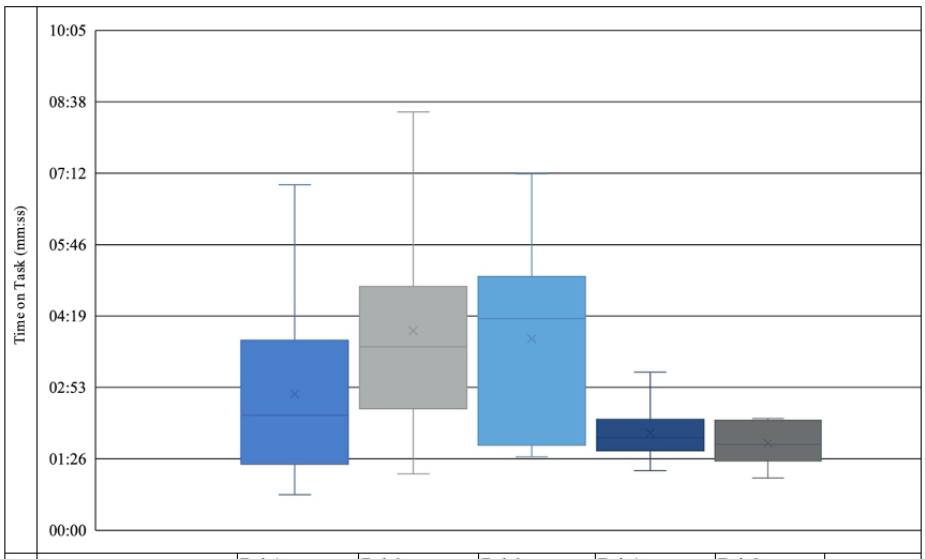


Table 4. Summary of usability issues grouped by topic.

Category and issue	Participants encountering issue, n (%)	Overall frequency	Individual severity ratings	Overall assessment of severity
Navigation				
Did not realize that the *Submit* button on the landing page was for moving forward in the survey	4 (57.1)	High	1, 3, 3, 4	Critical, high impact
Unclear how to enter first survey	1 (14.3)	Low	1	Critical, high impact
Clicking the back button resulted in an error message	2 (28.6)	Moderate	1, 2	Critical, high impact
Unclear how to return to the intervention after logging out	1 (14.3)	Low	2	Noncritical, moderate impact
Intervention features				
Small browser window caused text to wrap, which was difficult to read	1 (14.3)	Low	3	Noncritical, low impact
Unable to locate health tips	1 (14.3)	Low	1	Critical, high impact
Help and documentation				
Unsure if FAQ[a] was the right place to look for help	1 (14.3)	Low	4	Inconvenient, lowest impact
Identified a different page as the FAQ	2 (28.6)	Moderate	3, 4	Noncritical, low impact
Would try to contact MyChart (a clinical system) for help	1 (14.3)	Low	1	Critical, high impact
Content				
Text-heavy or wordy pages	4 (57.1)	High	4, 4, 4, 4	Inconvenient, lowest impact
Discomfort with psychological questions	5 (71.4)	High	3, 3, 4, 4, 4	Noncritical, low impact

[a]FAQ: frequently asked question.

Debriefing interviews revealed generally positive first impressions of the interface, and these positive impressions persisted throughout the test. Participants' dislikes and recommended changes were aligned with the usability issues documented during testing. When asked to comment on their first impressions of the intervention, 5 participants reported positive first impressions, of whom, 3 participants focused specifically on the visual impression of the interface and noted that it "look[ed] clean," appeared "pretty straightforward," and was "nice, clear, easy to read." Two participants reported negative first impressions and focused on the text-heavy nature of the homepage. Three participants noted that the *Submit* button made entry to the intervention somewhat confusing. Participants all reported that their first impressions did not substantially change as they navigated the intervention.

When asked what they liked best about the intervention, 5 participants reported that the resources and content were "informative," "useful," and "helpful," and 4 participants spoke to the design of the intervention and said that they appreciated how the content was "succinct," "clearly written," "not overwhelming," and that the fonts and color schemes were "pleasant." When asked what they liked least about the intervention, 3 participants reported that there was nothing they did not like, 3 participants did not like the wordiness, and 1 participant described it as a "general sense of clutter." One participant would have liked "prettier colors," and 2 participants did not like the questions that the interface posed about their voices: one participant wanted to know who would read and respond to her responses to these standard questions, and the other participant did not look forward to having to "write a lot of stuff out about my voice...that's not enticing."

When asked what they would change about the intervention, 2 participants said that there was nothing they would change. The other 5 participants recommended changes such as adding video content; providing short summaries of the content on each page; less text, especially on the front page; changing the *Submit* button on the landing page; and incorporating more consistent branding, such that the intervention would be more clearly connected to the clinic.

In participants' responses to the product review card, 89% of reaction words were positive. Only 1 participant chose negative words. Despite posttask ratings and observed issues that reflected more difficulty than we expected (Figure 4 and Table 4), participants' word choices on the product reaction card indicated positive feelings about the intervention. Out of 7 participants, 5 participants chose the word "organized," and 4 participants also chose "convenient" and "easy to use." Other common choices were "helpful," "relevant," and "useful." Although, it may seem that product reaction card results conflict with posttask ratings and observed issues, results from the debriefing interviews supported an overall positive participant reaction to the intervention. Participants were highly engaged

with the content itself and liked that it was web-based, although that was not the target of usability evaluations.

Overall, although participants experienced some challenges when completing tasks within the intervention, all usability metrics met or exceeded our predetermined goals regarding task completion, posttask ratings, and positivity of intervention descriptive words.

Discussion

Principal Findings

Our heuristic evaluation and usability test results provide evidence that REDCap is a useful platform for patient-facing web-based information and intervention delivery. Our findings demonstrate that the REDCap intervention was functional and usable for participants. Participants' comments demonstrated that they found the intervention and, by extension, its REDCap interface, to be one that they could imagine layering into their existing medical routines. Because the use of the REDCap platform for direct interaction in a patient-facing health intervention is relatively novel, and to the best of our knowledge, this is the first formal usability assessment of REDCap used in this manner, below we delineate specific recommendations for researchers wishing to develop REDCap for delivery of patient-facing interventions.

REDCap Development Recommendations

We found that the successful conversion of our initial custom website to a patient-facing web-based REDCap intervention aligned with and supported the overall user-friendliness of REDCap for investigator use. Our novel use of REDCap was further supported by the overall high acceptability and usability observed in our formal testing. We did encounter specific challenges, and additional information on strategies that proved useful for our team is offered below for others considering similar approaches to patient-facing information or intervention delivery. Some of these challenges and mitigating strategies may become obsolete as REDCap functionality is continually updated. For example, in a recent update, a data dictionary was created for ASIs, which greatly improved the ease of intervention development and troubleshooting. In addition, REDCap is developing options such as Mobile App and MyCap for use on mobile devices [32].

Our study also highlighted specific and persistent features of the REDCap intervention that detracted from user experience. Navigation challenges identified in heuristic evaluations—moving between help and survey queue pages, and finding "emergency exits" from the modules—suggest that REDCap is currently best suited to interactions that do not require extensive navigation between different modules. Additionally, while REDCap's error messaging was clear and timely, there were fewer options for making help documentation readily available throughout the intervention without a workaround such as a link menu at the bottom of the page. When we evaluated the issues that participants experienced when completing the usability test tasks, we found that some issues were related to the structure and limitations of REDCap as an interface. Thus, our usability findings also underscore the value of insights gained from research with end users and could help other researchers deploy REDCap as a patient-facing intervention delivery method.

We encountered a few issues in converting a custom website to a REDCap format; some issues remained unresolved and were tested in the usability evaluations. Issues highlighted by usability testing fell into 3 categories: conceptual expectations of website, nonintuitive navigation, and confusing site architecture (Table 5).

Table 5. Adaptations in REDCap to address heuristic and usability findings.

Challenges with REDCap[a] intervention delivery	Sample adaptation to enhance usability
Mismatch between format and the conceptual expectations of a website	
Heuristic analysis recommended better distribution of white space by moving information to the footer, header, or side menus where possible.	Participant resources links and survey page instructions were moved to the Survey Footer to separate them from module related text.
Participants recommended more branding visibility, as they appreciated the affiliation of the project with their clinic.	Aesthetics were constrained by limited options where logos could be added; combined logos were created to allow multiple entities to be represented.
Participants found page titles confusing and recommended clearer instructions and wording about the intervention and module titles.	Headers and text were revised and simplified to clarify instructions about the study and intervention and make the intervention consistently identifiable on each page.
Nonintuitive navigation through the program	
No independent home page functionality besides using the *Survey Queue* as a starting point, which participants found unfamiliar and confusing.	A site map was not possible within REDCap. Therefore, study status graphics were added to the first and last page of each module to show the participant's progression through the intervention.
Participants struggled to tell how far along they were in the program, as the survey queue did not show what was forthcoming when using *Automatic Survey Invitations*.	Page numbers were added to show progression through each module.
Participants found that saving and returning using the randomly generated code for re-entry was nonintuitive and easy to miss when leaving a survey, making returning to the intervention difficult.	The *Survey Login* was enabled to use participant email to log into REDCap instead of a random generated code. Instructions for navigation in the FAQ[b] were added and linked to the FAQ in the *Automatic Survey Invitation* email(s).
Participants experienced difficulty returning to REDCap intervention pages after clicking on a hyperlink due lack of ability to link back to other instruments within a survey.	The number of embedded hyperlinks was minimized. Where hyperlinks were unavoidable, instructions were added, eg, how to navigate back to the next part of the intervention from the patient resources webpage.
Confusing site architecture	
REDCap's participant-facing interface was the survey format, and participants struggled with hardcoded survey labels and buttons such as *Survey Login* or *Close Survey*.	Instructions were revised to say "survey" instead of "assessment" or "questionnaire." When removing survey labels was not possible, such as for instructions in the linked tips and help documentation, descriptive text with instructions was added, eg, "Click 'Close Survey' to close this window. Then go back to the program page."
Using the *Survey Queue* as the home page for the intervention confused some participants because the program was not a survey in the typical sense.	Visible use of *Survey Queue* was replaced with study status graphics at the beginning and end of each module to limit the amount of "survey" titles and buttons. The *Stealth Queue* external plug-in was used to prevent the survey queue from automatically displaying at the end of a survey.
To advance, participants needed to click the *Submit* button, even if nothing was being submitted, such as after viewing educational material.	The number of instruments per module was reduced to limit the number of *Submit* buttons.

[a]REDCap-specific terms are in italics.

[b]FAQ: frequently asked question.

Strengths and Limitations

Key strengths of this study were the multidimensional nature of our assessment, with both heuristic evaluation and formal usability testing. For the latter, we incorporated objective and subjective task-based data, such as timing and scoring, as well as open-ended data formats, such as think-aloud responses and a debriefing interview. This layered structure allowed for a rich examination of multiple types of usability data from the patient-facing REDCap intervention. Another key strength of the study is that the team was multidisciplinary and included expertise in usability, writing studies, engineering, psychology, voice, and medicine, which allowed the incorporation of perspectives from multiple areas, which in turn, strengthened the potential generalizability of findings. In addition, the study was completed with patients from the target population for the intervention, which increased the face and content validity of our findings [33]. Perhaps the greatest strength of this study is that it offers a practical approach to a challenging problem: how to translate helpful content into a format that is usable for patient participants in an affordable, transparent manner. Our findings allow for ready expansion to create comparison arms for our existing studies and would be useful to other research teams pursuing similar avenues of investigation and to clinicians or others who may wish to deliver information to patients and clients in an interactive secure manner. The limitations of the study include its small sample size and the limited diversity therein, which both impact generalizability. However, the

strengths of the study outweigh its limitations, and we hope to address these limitations in future studies.

Considerations for Researchers

REDCap is an appealing platform for a web-based intervention because of its ease of use for researchers and participants, favorable cost and accessibility, and overall effective usability. Furthermore, REDCap is an evolving resource, with additional functionalities frequently being added. In some cases, new functionalities alter the behavior of current active projects, and in other cases, new functionalities offer helpful solutions to important challenges. We recommend that researchers developing an intervention with REDCap's current capabilities consider customizing REDCap delivery based on intervention needs using tools such as field variables, structured module timing, and piping options; minimizing the use of tools that display the hardcoded term "survey" in the text, such as survey queue, submit survey buttons (unless the study is purely a survey); enabling survey log-in and provide clear information on how to navigate in and out of the intervention; making stage of progression through the intervention clear (eg, page numbers, study status graphic); and paying close attention to REDCap updates that may change functionality.

REDCap may be particularly helpful for developing functional intervention prototypes, because it allows researchers to efficiently incorporate changes based on participant feedback for rapid testing of iterations of the content and format. It also allows for the efficient creation of comparison study arms. Interventions developed in this manner could be optimized and permanently used in REDCap, or used as a functional prototype or model for a custom website. Overall, our findings suggest that REDCap can effectively be used for patient-facing intervention delivery, particularly with adaptations such as those suggested above to optimize its usability. We anticipate that, as REDCap evolves and continues to partner with clinicians and researchers, its applicability will expand even further, reducing barriers for teams offering patient-facing interventions.

Acknowledgments

This study was supported by funding from the National Institute on Deafness and other Communication Disorders (K23DC016335) and the National Center for Advancing Translational Sciences (KL2TR000113 and UL1TR002494) of the National Institutes of Health, and a Clinical Scientist Development Award (SM) from the American College of Surgeons and the Triological Society. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health or the American College of Surgeons.

Conflicts of Interest

SM is supported by research funding from the National Institutes of Health and the American College of Surgeons/Triological Society. The content of this publication does not reflect the official views of either the National Institutes of Health or the American College of Surgeons/Triological Society.

Multimedia Appendix 1
Usability test tasks.
[DOCX File , 22 KB - humanfactors_v9i1e26461_app1.docx]

Multimedia Appendix 2
Debriefing interview questions.
[DOCX File , 22 KB - humanfactors_v9i1e26461_app2.docx]

References

1. Cohen SM. Self-reported impact of dysphonia in a primary care population: an epidemiological study. Laryngoscope 2010 Oct;120(10):2022-2032. [doi: 10.1002/lary.21058] [Medline: 20830762]
2. Schwartz SR, Cohen SM, Dailey SH, Rosenfeld RM, Deutsch ES, Gillespie MB, et al. Clinical practice guideline: hoarseness (dysphonia). Otolaryngol Head Neck Surg 2009 Sep;141(3 Suppl 2):S1-S31. [doi: 10.1016/j.otohns.2009.06.744] [Medline: 19729111]
3. Titze IR, Lemke J, Montequin D. Populations in the U.S. workforce who rely on voice as a primary tool of trade: a preliminary report. J Voice 1997 Sep;11(3):254-259. [doi: 10.1016/s0892-1997(97)80002-1] [Medline: 9297668]
4. Cohen SM, Kim J, Roy N, Asche C, Courey M. Direct health care costs of laryngeal diseases and disorders. Laryngoscope 2012 Jul;122(7):1582-1588. [doi: 10.1002/lary.23189] [Medline: 22544473]
5. Misono S, Meredith L, Peterson CB, Frazier PA. New perspective on psychosocial distress in patients with dysphonia: the moderating role of perceived control. J Voice 2016 Mar;30(2):172-176. [doi: 10.1016/j.jvoice.2015.02.002] [Medline: 25795347]
6. Frazier P, Meredith L, Greer C, Paulsen JA, Howard K, Dietz LR, et al. Randomized controlled trial evaluating the effectiveness of a web-based stress management program among community college students. Anxiety Stress Coping 2015;28(5):576-586. [doi: 10.1080/10615806.2014.987666] [Medline: 25420030]

7. Hintz S, Frazier PA, Meredith L. Evaluating an online stress management intervention for college students. J Couns Psychol 2015 Apr;62(2):137-147. [doi: 10.1037/cou0000014] [Medline: 24635586]
8. Atlantis E, Fahey P, Foster J. Collaborative care for comorbid depression and diabetes: a systematic review and meta-analysis. BMJ Open 2014 Apr 12;4(4):e004706 [FREE Full text] [doi: 10.1136/bmjopen-2013-004706] [Medline: 24727428]
9. Huang Y, Wei X, Wu T, Chen R, Guo A. Collaborative care for patients with depression and diabetes mellitus: a systematic review and meta-analysis. BMC Psychiatry 2013 Oct 14;13:260 [FREE Full text] [doi: 10.1186/1471-244X-13-260] [Medline: 24125027]
10. Olajos-Clow J, Costello E, Lougheed MD. Perceived control and quality of life in asthma: impact of asthma education. J Asthma 2005 Nov;42(9):751-756. [doi: 10.1080/02770900500308080] [Medline: 16316869]
11. Carney RM, Freedland KE, Stein PK, Skala JA, Hoffman P, Jaffe AS. Change in heart rate and heart rate variability during treatment for depression in patients with coronary heart disease. Psychosom Med 2000;62(5):639-647. [doi: 10.1097/00006842-200009000-00007] [Medline: 11020093]
12. Doering LV, McGuire A, Eastwood J, Chen B, Bodán RC, Czer LS, et al. Cognitive behavioral therapy for depression improves pain and perceived control in cardiac surgery patients. Eur J Cardiovasc Nurs 2016 Oct 09;15(6):417-424 [FREE Full text] [doi: 10.1177/1474515115592292] [Medline: 26115954]
13. Hill LK, Sims Wright R, Aiken-Morgan AT, Gamaldo A, Edwards CL, Whitfield KE. Perceived control predicts pulse pressure in African American men: the Baltimore study of Black aging. Ethn Dis 2015 Aug 07;25(3):263-270 [FREE Full text] [doi: 10.18865/ed.25.3.263] [Medline: 26676156]
14. Nguyen-Feng VN, Frazier PA, Stockness A, Narayanan A, Merians AN, Misono S. Web-based perceived present control intervention for voice disorders: a pilot study. J Voice 2020 Mar;34(2):300.e1-300.e9 [FREE Full text] [doi: 10.1016/j.jvoice.2018.08.006] [Medline: 30227981]
15. 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]
16. Shelley AJ, McDonald KA, McEvoy A, Sauder M, Kanigsberg N, Zemek R, et al. Usability, satisfaction, and usefulness of an illustrated eczema action plan. J Cutan Med Surg 2018 Jul 13;22(6):577-582. [doi: 10.1177/1203475418789028] [Medline: 30003800]
17. Purkayastha S, Addepally SA, Bucher S. Engagement and usability of a cognitive behavioral therapy mobile app compared with web-based cognitive behavioral therapy among college students: randomized heuristic trial. JMIR Hum Factors 2020 Feb 03;7(1):e14146 [FREE Full text] [doi: 10.2196/14146] [Medline: 32012043]
18. Beatty AL, Magnusson SL, Fortney JC, Sayre GG, Whooley MA. VA FitHeart, a mobile app for cardiac rehabilitation: usability study. JMIR Hum Factors 2018 Jan 15;5(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.8017] [Medline: 29335235]
19. Chung S, Hacker ED, Rawl S, Ellis R, Bakas T, Jones J, et al. Using Facebook in recruiting kidney transplant recipients for a REDCap study. West J Nurs Res 2019 Mar 05:193945919832600. [doi: 10.1177/0193945919832600] [Medline: 30836840]
20. Cochran AR, Raub KM, Murphy KJ, Iannitti DA, Vrochides D. Novel use of REDCap to develop an advanced platform to display predictive analytics and track compliance with enhanced recovery after surgery for pancreaticoduodenectomy. Int J Med Inform 2018 Nov;119:54-60. [doi: 10.1016/j.ijmedinf.2018.09.001] [Medline: 30342686]
21. Crane S, Comer RS, Arenson AD, Draucker C. Using REDCap to facilitate web-based therapeutic intervention research. Nurs Res 2019;68(6):483-487. [doi: 10.1097/nnr.0000000000000367]
22. Corkum PV, Reid GJ, Hall WA, Godbout R, Stremler R, Weiss SK, et al. Evaluation of an internet-based behavioral intervention to improve psychosocial health outcomes in children with insomnia (better nights, better days): protocol for a randomized controlled trial. JMIR Res Protoc 2018 Mar 26;7(3):e76 [FREE Full text] [doi: 10.2196/resprot.8348] [Medline: 29581089]
23. Mayberry LS, Berg CA, Harper KJ, Osborn CY. The design, usability, and feasibility of a family-focused diabetes self-care support mHealth intervention for diverse, low-income adults with type 2 diabetes. J Diabetes Res 2016;2016:7586385 [FREE Full text] [doi: 10.1155/2016/7586385] [Medline: 27891524]
24. Fuller TE, Pong DD, Piniella N, Pardo M, Bessa N, Yoon C, et al. Interactive digital health tools to engage patients and caregivers in discharge preparation: implementation study. J Med Internet Res 2020 Apr 28;22(4):e15573 [FREE Full text] [doi: 10.2196/15573] [Medline: 32343248]
25. Barnum C. Usability Testing Essentials: Ready, Set ...Test! (2nd Ed). Amsterdam: Morgan Kaufmann; 2020.
26. Nielsen J. 10 usability heuristics for user interface design. Nielsen Norman Group. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2020-11-13]
27. Kessler M, Breuch L, Stambler D, Campeau K, Riggins O, Feddema E, et al. J Technical Writing and Comm 2021 Oct 19;51(4):380-406. [doi: 10.1177/00472816211044498]
28. Arffa RE, Krishna P, Gartner-Schmidt J, Rosen CA. Normative values for the voice handicap index-10. J Voice 2012 Jul;26(4):462-465. [doi: 10.1016/j.jvoice.2011.04.006] [Medline: 21816570]

29. Benedek J, Miner T. Measuring desirability: new methods for evaluating desirability in a usability lab setting. 2002 Presented at: Usability Professionals Association; July 8-12, 2002; Orlando, Florida.
30. Moran K. Using the Microsoft desirability toolkit to test visual appeal. Nielsen Norman Group. 2016 Feb 28. URL: https://www.nngroup.com/articles/microsoft-desirability-toolkit/ [accessed 2021-05-05]
31. Plain language guidelines. Plain Language Action and Information Network. URL: https://www.plainlanguage.gov/guidelines/ [accessed 2022-03-15]
32. REDCap. Projectredcap. URL: https://projectredcap.org/software/mobile-app/ [accessed 2020-11-13]
33. Connell J, Carlton J, Grundy A, Taylor Buck E, Keetharuth AD, Ricketts T, et al. The importance of content and face validity in instrument development: lessons learnt from service users when developing the recovering quality of life measure (ReQoL). Qual Life Res 2018 Jul;27(7):1893-1902 [FREE Full text] [doi: 10.1007/s11136-018-1847-y] [Medline: 29675691]

Abbreviations

ASI: Automatic Survey Invitation
FAQ: frequently asked question
HIPAA: Health Insurance Portability and Accountability Act
REDCap: Research Electronic Data Capture

Edited by A Kushniruk; submitted 12.12.20; peer-reviewed by C Barnum, W Martinez, E Portales-Casamar; comments to author 15.03.21; revised version received 10.05.21; accepted 08.12.21; published 25.03.22.

Please cite as:
Stambler DM, Feddema E, Riggins O, Campeau K, Breuch LAK, Kessler MM, Misono S
REDCap Delivery of a Web-Based Intervention for Patients With Voice Disorders: Usability Study
JMIR Hum Factors 2022;9(1):e26461
URL: https://humanfactors.jmir.org/2022/1/e26461
doi:10.2196/26461
PMID:35333191

Original Paper

Designing a Novel Clinician Decision Support Tool for the Management of Acute Diarrhea in Bangladesh: Formative Qualitative Study

Rochelle K Rosen[1], PhD; Stephanie C Garbern[2], MPH, MD; Monique Gainey[3], MSc, MPH; Ryan Lantini[1], MA; Sabiha Nasrin[4], MBBS, MPH; Eric J Nelson[5], MD, PhD; Nour Elshabassi[6], MPH; Nur H Alam[4], MBBS, MD, PhD; Sufia Sultana[4], MA; Tahmida Hasnin[4], BA; Kexin Qu[7], MSc; Christopher H Schmid[7], PhD; Adam C Levine[2], MPH, MD

[1]Center for Behavioral and Preventive Medicine, The Miriam Hospital, Providence, RI, United States

[2]Department of Emergency Medicine, Warren Alpert Medical School, Brown University, Providence, RI, United States

[3]Department of Emergency Medicine, Rhode Island Hospital, Providence, RI, United States

[4]Nutrition and Clinical Services Division, International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh

[5]Emerging Pathogens Institute, University of Florida, Gainesville, FL, United States

[6]Brown University, Providence, RI, United States

[7]Department of Biostatistics, School of Public Health, Brown University, Providence, RI, United States

Corresponding Author:
Rochelle K Rosen, PhD
Center for Behavioral and Preventive Medicine
The Miriam Hospital
164 Summit St
Coro Building Suite 134
Providence, RI, 02906
United States
Phone: 1 401 793 8182
Email: rrosen@lifespan.org

Abstract

Background: The availability of mobile clinical decision support (CDS) tools has grown substantially with the increased prevalence of smartphone devices and apps. Although health care providers express interest in integrating mobile health (mHealth) technologies into their clinical settings, concerns have been raised, including perceived disagreements between information provided by mobile CDS tools and standard guidelines. Despite their potential to transform health care delivery, there remains limited literature on the provider's perspective on the clinical utility of mobile CDS tools for improving patient outcomes, especially in low- and middle-income countries.

Objective: This study aims to describe providers' perceptions about the utility of a mobile CDS tool accessed via a smartphone app for diarrhea management in Bangladesh. In addition, feedback was collected on the preliminary components of the mobile CDS tool to address clinicians' concerns and incorporate their preferences.

Methods: From November to December 2020, qualitative data were gathered through 8 web-based focus group discussions with physicians and nurses from 3 Bangladeshi hospitals. Each discussion was conducted in the local language—Bangla—and audio recorded for transcription and translation by the local research team. Transcripts and codes were entered into NVivo (version 12; QSR International), and applied thematic analysis was used to identify themes that explore the clinical utility of an mHealth app for assessing dehydration severity in patients with acute diarrhea. Summaries of concepts and themes were generated from reviews of the aggregated coded data; thematic memos were written and used for the final analysis.

Results: Of the 27 focus group participants, 14 (52%) were nurses and 13 (48%) were physicians; 15 (56%) worked at a diarrhea specialty hospital and 12 (44%) worked in government district or subdistrict hospitals. Participants' experience in their current position ranged from 2 to 14 years, with an average of 10.3 (SD 9.0) years. Key themes from the qualitative data analysis included current experience with CDS, overall perception of the app's utility and its potential role in clinical care, barriers to and facilitators of app use, considerations of overtreatment and undertreatment, and guidelines for the app's clinical recommendations. Participants

felt that the tool would initially take time to use, but once learned, it could be useful during epidemic cholera. Some felt that clinical experience remains an important part of treatment that can be supplemented, but not replaced, by a CDS tool. In addition, diagnostic information, including mid-upper arm circumference and blood pressure, might not be available to directly inform programming decisions.

Conclusions: Participants were positive about the mHealth app and its potential to inform diarrhea management. They provided detailed feedback, which developers used to revise the mobile CDS tool. These formative qualitative data provided timely and relevant feedback to improve the utility of a CDS tool for diarrhea treatment in Bangladesh.

(JMIR Hum Factors 2022;9(1):e33325) doi:10.2196/33325

KEYWORDS

clinical decision support tools; diarrhea management; focus group; formative qualitative research; low- and middle-income countries; mobile phone

Introduction

Background

Mobile technology has had a major impact on the rapid access and transfer of information globally. Today, it is estimated that >5 billion people have mobile devices, with over half of these being smartphones [1]. In the health care sector, smartphones are increasingly used to improve communication between physicians and patients as well as to improve clinical decision-making. With >300,000 mobile health (mHealth) apps available in major app stores, the availability of mobile clinical decision support (CDS) to health care professionals has grown substantially with the increased prevalence of smartphone devices and apps [2,3]. Defined as information systems designed to improve clinical decision-making, traditional forms of CDS range from integration of electronic health records to software apps providing guidelines on a clinical topic [4]. A survey conducted in the United States by the Health Information and Management Systems Society [5] revealed that nearly 90% of providers use mobile devices to engage with patients, whereas a US-based survey analyzing physician information sources [6] found that 72% of physicians use a smartphone or tablet to access drug information and 63% to access medical research.

Although existing CDS systems enable health care professionals to leverage the benefits of technology and information for their clinical practice, many individual, institutional, and technological barriers affect the engagement of clinicians with these new technologies. For instance, in a study conducted in the United Kingdom, physicians found it difficult to integrate mobile CDS into their pattern of work, prompting them to seek alternative sources of CDS [7]. Several studies examining mobile CDS use by physicians in the United Kingdom and United States found that the uptake of mobile CDS was hindered because of their perception that using or adopting such technology included having to choose between suggestions given by mobile CDS and traditionally *trusted* sources, disagreement between information provided by mobile CDS and standard guidelines, and the belief that the use of mobile CDS would be perceived as being unprofessional by the patient [7,8].

Such concerns are not limited to high-income countries (HICs). A cross-sectional study aimed at assessing smartphone medical app use among physicians in Ethiopia found that the perceived usefulness of the app was one of the most notable factors associated with medical app use by physicians along with attitude, internet access, technical skills, and information technology support staff [9]. Furthermore, providers in low- and middle-income countries (LMICs) have expressed willingness to use mHealth tools and perceive such technology as playing an important role in reducing health care barriers [9-11]. A comparative study analyzing the limitations of mobile CDS app adoption and use by clinicians in LMICs versus HICs found that users from LMICs, primarily those who practice on their own in rural settings, used the app more frequently and rated the app as more important for their practice [11]. However, another study has shown that the use of CDS in resource-limited settings was associated with stronger adherence to standard guidelines. More specifically, a randomized controlled trial in Bangladesh found that electronic decision support improved treatment changes that were more consistent with the World Health Organization (WHO) guidelines [12]. However, even with such successes, engagement and implementation of such tools are hindered by limited awareness of mHealth, illiteracy, variable quality of care, and poor network connectivity [10]. In both HICs and LMICs, although providers express interest in integrating mHealth technologies in a clinical setting and have reported positive perceptions toward using mobile CDS as part of everyday practice, many are unconvinced of its overall clinical usefulness for patient outcomes because of the lack of literature on this topic [7,13-15].

Objectives

Despite the potential to transform health care delivery, much still remains unknown about the clinical usability of mobile CDS from a provider perspective, especially in LMICs. As such, the aim of this study is to describe providers' perceptions about the utility of a mobile CDS tool that integrates predictive models for dehydration assessment in patients with acute diarrhea in Bangladesh and to seek their feedback on the preliminary components of this CDS tool. Qualitative data were gathered through focus groups, with physicians and nurses working in diverse clinical settings, including specialty research and general public hospitals. In consulting with clinicians, we hope to better adapt, and increase user confidence in, the predictive models and treatment recommendations provided by this tool. By seeking feedback on the tool's layout and design to ensure that it fits appropriately into different clinical contexts, this formative

qualitative research better enables the designers to build a CDS tool that anticipates and addresses the aforementioned barriers.

Methods

Study Design and Setting

Qualitative data were collected in a series of focus group discussions (FGDs) from November to December 2020 among clinicians working at 3 distinct hospitals in Bangladesh as part of the *Novel, Innovative Research for Understanding Dehydration in Adults and Kids* (NIRUDAK; which also means *dehydration* in Bangla) study. NIRUDAK is an ongoing research effort to develop diagnostic models and incorporate them in a mobile app to support clinical decisions in the treatment and assessment of dehydration severity in patients with acute diarrhea. The focus groups obtained feedback from nurses and physicians on the clinical utility of the NIRUDAK mHealth app (NIRUDAK app) to understand the current use of mHealth and other CDS tools, understand factors clinicians consider essential in treating patients with diarrhea, review the preliminary app design and content, and seek feedback on app development before a pilot test and trial of its clinical use.

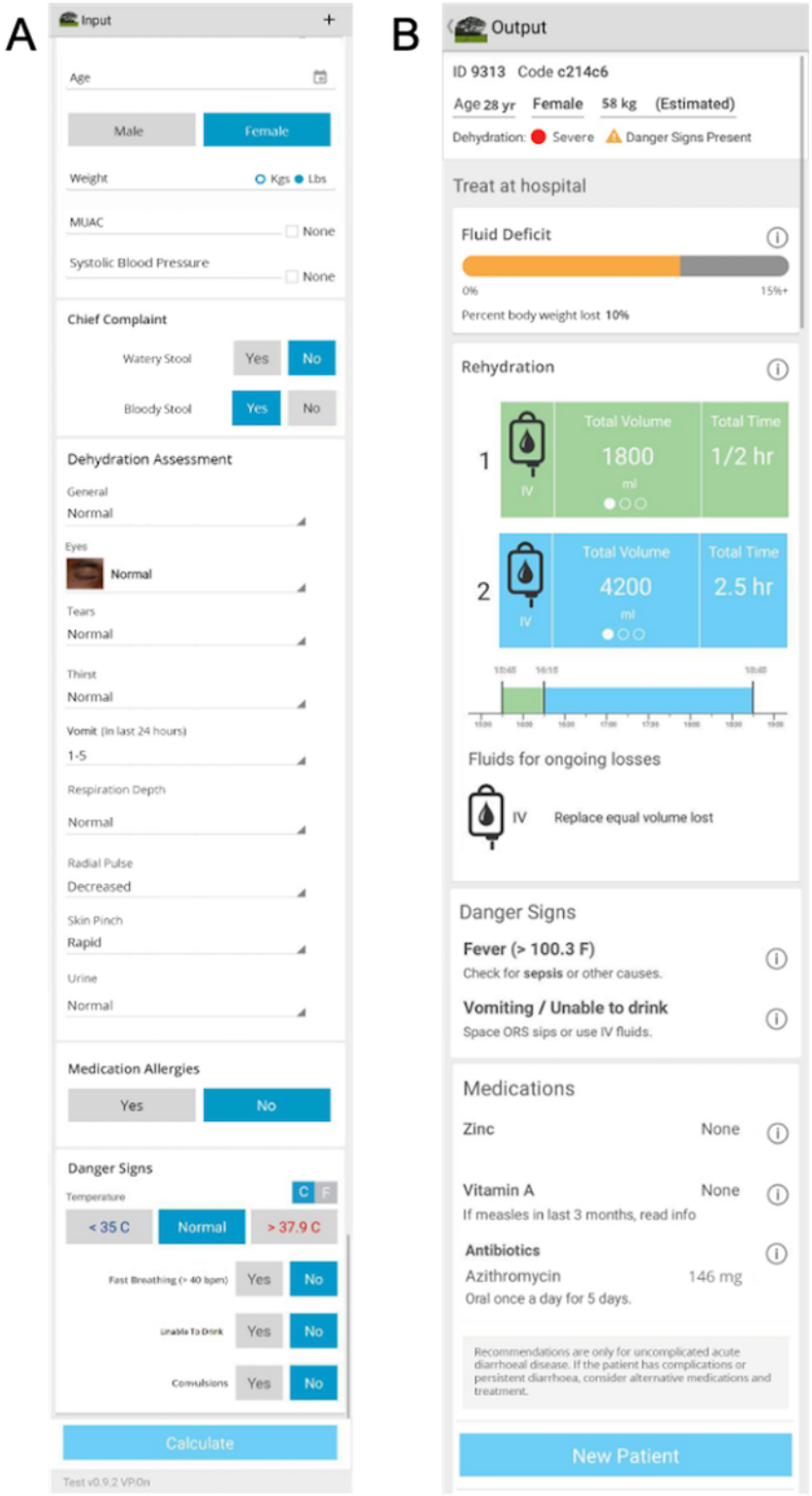
Owing to travel restrictions imposed by the COVID-19 pandemic, all data were gathered using a web-based platform (Zoom; Zoom Video Communications, Inc). Data were collected from clinicians working at three hospitals in Bangladesh: (1) the International Centre for Diarrhoeal Disease Research, Bangladesh's (icddr,b) Dhaka Hospital; (2) Narayanganj General Victoria District Hospital; and (3) Shaheed Ahsan Ullah Master General Hospital (also known as Tongi Upazilla or Subdistrict Hospital). icddr,b's Dhaka Hospital is a 350-bed, not-for-profit international research hospital specializing in the treatment of diarrheal illnesses and providing clinical services at no charge to over 100,000 patients with acute diarrhea a year from a catchment area of over 17 million people from the city of Dhaka and its nearby rural districts [16]. Narayanganj General Victoria District Hospital is a 100-bed facility in the town of Narayanganj [17]. Treating 30 to 40 patients with diarrhea per day, this district hospital works as a referral center to primary-level facilities, such as Tongi Upazilla, and is also a site of the nationwide diarrhea surveillance program run by the Institute of Epidemiology, Disease Control and Research and icddr,b. The Tongi Upazilla Hospital, a 250-bed hospital, acts as a primary-level health facility in the district of Gazipur [17].

Dehydration Management and the NIRUDAK App

Appropriate rehydration with oral and intravenous fluids is the most important treatment for acute diarrhea and requires an accurate assessment of dehydration level [18-24]. Patients with mild to moderate dehydration can be treated with oral rehydration solution in the outpatient setting, whereas those without any dehydration often times need only instructions for management at home [23-25]. Patients with severe dehydration require intravenous fluids in a hospital setting to avoid hemodynamic instability, organ ischemia, and death [23-25]. As the severity of illness can vary greatly among patients, accurate assessment of dehydration status remains a critical step in diarrhea management and can reduce morbidity and mortality that results from both overhydration and underhydration of patients [23-25].

The NIRUDAK app was developed to incorporate several clinical diagnostic models, derived using logistical regression, for assessment of dehydration severity in patients with acute diarrhea aged >5 years (full and simplified NIRUDAK models) and in children aged <5 years (DHAKA [Dehydration: Assessing Kids Accurately] score) [23,26]. On the basis of a review of literature and consultation with expert clinicians at icddr,b, a total of 18 signs or symptoms of dehydration were selected to derive the full NIRUDAK model. A total of 11 more basic clinical predictors were selected for the simplified NIRUDAK model with the intention that it could be used in settings where resources may be limited (ie, places without the ability to measure blood pressure, which is required for the full model) [23]. After assessing each model's performance, the final full NIRUDAK model included 8 predictors of dehydration, and the final simplified NIRUDAK model included 7 (Multimedia Appendix 1) [23]. The algorithms of both models were then incorporated into a mobile app prototype. The prototype was derived from an mHealth CDS (*Rehydration Calculator*) that adapted paper-based WHO guidelines to the digital medium [27,28]. The prototype allowed for clinicians to enter a patient's symptoms in the input screen (Figure 1A) and to receive the patient's dehydration severity level and specific treatment recommendations on the output screen (Figure 1B). Once validated, the NIRUDAK app will enable dehydration severity level assessment (none, some, or severe) and improve the management of patients with acute diarrhea in low-resource settings.

Figure 1. Screenshots of the NIRUDAK app's input and output screens shown to the participants during focus group discussions. (A) The input screen illustrates where clinicians enter relevant data based on a clinical assessment. (B) The output screen displays the patient's dehydration severity level and fluid deficit as well as targeted recommendations for rehydration. NIRUDAK: Novel, Innovative Research for Understanding Dehydration in Adults and Kids.



Study Participants

A total of 8 focus groups were conducted. Of the 8 focus groups, 2 (25%) were conducted with clinicians from each of the district and subdistrict hospitals, one with nurses and the other with physicians. Moreover, of the 8 focus groups, 4 (50%) were conducted with icddr,b providers, 2 (25%) with physicians and 2 (25%) with nurses. Each focus group ranged from 2 to 4 participants, with most focus groups being conducted with 4 participants. The number of participants per focus group was deliberately kept low, in keeping with best practices for remote focus groups [29,30].

Data Collection

All FGDs were conducted using a web-based platform (Zoom) in Bangla and facilitated by a member of the Bangladesh-based research team. Facilitators used a written focus group agenda, which ensured that all groups were facilitated similarly, and all participants were asked the same series of questions. The agenda asked about the current use of mHealth tools and then presented a standardized case of a patient with diarrhea. Clinicians were asked to identify the patient data essential to determining diarrheal treatment, including rehydration. Facilitators then showed a short video of a prototype app, demonstrating its key features, input screens, and components. Still images of each screenshot were then shown; participants were asked for feedback on input and output screens and to choose between different possible layouts and models for the final mHealth app. Most of the focus groups were approximately 1-hour long, and typed transcripts ranged from 26 to 29 pages. Each discussion was audio recorded for transcription and translation by the local research team, which included the facilitators.

Transcription and translation were conducted in multiple steps and took approximately 8 weeks to complete. First, audio recordings were transcribed into Bangla. Next, another team member reviewed the audio and Bangla transcript to ensure accuracy and that all data had been deidentified. The Bangla transcripts were then translated into English by a research team member proficient in written and spoken English. The English transcripts were reviewed in Bangladesh by a third team member for accuracy. Finally, a US-based research team member read each English transcript to determine if any further clarification was needed. The Bangladesh-based research team reviewed and resolved all translation clarification requests. Once these were addressed, the English transcripts were considered finalized and used in the coding and data analysis process.

Data Analysis

The research team used applied thematic analysis, a rigorous yet inductive approach designed to identify and examine themes from textual data [31]. Several steps were conducted to augment the rigor and credibility of the qualitative analysis, with coding occurring in 2 major stages.

In stage 1, coding structures were derived inductively as themes and repetitions emerged from reading the first 3 transcripts line by line and systematically categorizing emergent codes. A codebook was created to index and define each emergent code. An audit trial was used to document the iterative process of consolidating and creating emergent codes. In stage 2, the 8 transcripts were independently coded by 2 analysts who then met to compare codes and resolve discrepancies to establish intercoder agreement. Transcripts and codes were entered into NVivo (version 12; QSR International) for analysis [32].

Next, summaries of concepts and themes were generated from reviews of the key aggregated coded data. Coding summaries report participant comments for relevant codes, tracking the number of comments and the distribution of the data among participants. Data in the code summaries were organized by clinician category (nurse or physician) and hospital type (specialty, district, or subdistrict) for easy comparison of similarities and differences among the participants in those categories. Thematic memos, which gathered data from several summaries into key topic areas, were then written and used for the final analysis.

Five qualitative team members participated in the analysis: three coders (the study project coordinator [MG]; a master's level analyst [RL]; and one of the coinvestigators, a PhD anthropologist [RKR]), who were supported by two coinvestigators with experience in treating diarrhea in Bangladesh (SCG and SN). All data and memos were interpreted in collaboration with both US- and Bangladesh-based research team members and with the principal investigator (ACL) and app designer (EJN; both doctors of medicine with extensive global health and diarrheal disease expertise) for the purpose of identifying themes that explore the clinical utility of an mHealth app to assess dehydration severity in patients aged >5 years with acute diarrhea.

Ethical Approval and Consent to Participate

Ethical approval for the formative phase of the NIRUDAK study was obtained from the icddr,b (PR-20048) and the Lifespan-Rhode Island Hospital (1624612) institutional review boards.

Results

Overview

In total, 8 focus groups were attended by 27 participants. Of these 27 participants, 14 (52%) were nurses and 13 (48%) were physicians; in addition, of the 27 participants, 15 (56%) worked at icddr,b Dhaka Hospital and 12 (44%) worked in government district and subdistrict hospitals. The participants' experience in their current position ranged from 2 to 14 years, with an average of 10.3 (SD 9.0) years. Additional demographic information is shown in Table 1.

Several key themes emerged from the qualitative data analysis: current experience with CDS, overall perception of the app's utility and its potential role in clinical care, feedback on specific app details, barriers to and facilitators of app use, considerations of overtreatment and undertreatment, and guidelines for the app's clinical recommendations. We present these themes in our results and consider the implications of each for the CDS tool development in the discussion that follows.

Table 1. Baseline demographics (N=27).

Characteristics	Values
Age (years), n (%)	
25-34	6 (22)
35-44	16 (59)
45-54	4 (14)
55-64	1 (3)
Sex, n (%)	
Men	9 (33)
Women	18 (66)
Position and degree, n (%)	
Nurse	14 (51)
Diploma	5 (35)
Bachelor's degree	4 (28)
Master's degree	5 (35)
Physician	13 (48)
MBBS[a]	8 (61)
Master's degree	5 (38)
Monthly household income[b], n (%)	
10,001-50,000 BDT (US $116-580)	7 (25)
50,001-100,000 BDT (US $581-1160)	6 (22)
>100,000 BDT (>US $1160)	14 (51)
Experience in current position (years), mean (SD)	10.3 (9.0)
Hospital location, n (%)	
icddr,b[c]	15 (55)
Tongi Upazilla Subdistrict Hospital	7 (25)
Narayanganj General Victoria District Hospital	5 (18)

[a]MBBS: Bachelor of Medicine, Bachelor of Surgery (a degree for physicians in Bangladesh).

[b]At the time of the focus group discussion, the US dollar to Bangladesh taka exchange rate was US $1=83.3 BDT.

[c]icddr,b: International Centre for Diarrhoeal Disease Research, Bangladesh.

Current Experience With CDS

Many, if not all, icddr,b participants (both physicians and nurses) indicated that they had previous experience with CDS and web-based tools. icddr,b clinicians had used SHEBA, an integrated, computerized, and paperless hospital information management system that has been in use since 2009 [33]. The clinical system is installed on all desktops and is the hospital system for patient records. Participants indicated that instructions related to patient follow-up and discharge can be entered into the system as well. District and subdistrict hospital clinicians had some CDS tool experience, including a diarrhea management tool; however, overall, they reported limited practice experience. Many participants, regardless of the setting, mentioned using the Bangladesh Drug Information Management System, a software-based information app used on a mobile phone. Some participants mentioned using UpToDate, an evidence-based clinical resource that includes several medical calculators. Finally, a few participants mentioned using apps for researching literature or for calculations, such as BMI.

Overall Perception of the App's Utility and Role in Clinical Care

Most participants were enthusiastic about the NIRUDAK app. Participants felt that it would take some time to learn to use the app but that once learned, the app would be easy to use and have the potential to ease or decrease their workload over time. A participant noted that when they started practicing medicine, patient assessment was tracked on paper and is now being tracked by computer. They noted the following about using this app:

> *[Using the app] is a matter of time, also learning. We want to assess [dehydration] with less things. By this I mean...having less buttons or features so our work*

> *will be easier. If that can happen, that will be good for us.* [specialty hospital nurse]
>
> *To start using a new thing, initially some problems will occur, which is normal. But it's an excellent app. If anyone uses [this app], it definitely will be beneficial, and it's easy. Not only doctors but also nurses can use it comfortably and, in many cases, our [work] load will decrease. Yes, [the app] is excellent, I think, it can be used.* [government hospital physician]

Confidence in the app was explicitly discussed in 2 FGDs with nurses. In these FGDs, the nurses reported enthusiasm for the app and felt confidence in use of the app: "You can send this app to any end of the world without hesitation" [specialty hospital nurse].

Several participants noted that the app and its recommendations should not exist in isolation from the clinician and could not replace clinician experience. For instance, participants noted that clinician decision-making for a patient with diarrhea may be more complex than solely determining the level of dehydration. Other comorbidities, such as diabetes and electrolyte imbalances, are important to recognize and may impact the management of diarrhea and dehydration versus the standardized output from the app. Several participants felt that clinical experience could be relevant to decide when the app use was appropriate; for example, clinicians must first make a determination that patients have dehydration versus sepsis in which fluid management strategies may differ substantially: "[If] I give the total [amount of] fluid [recommended by the app] for a patient with severe dehydration and [the patient also has] sepsis, in that case it will be detrimental for the patient" [specialty hospital physician].

Feedback on Specific App Details

Overview

During the focus group, participants reviewed screenshots of the NIRUDAK app prototype, including images of the input and output screens. Here, we present comments from our participants about the following three specific app elements: age, danger signs, and the fluid deficit bar used in treatment recommendations. Each element is described in the next sections, with associated participant comments. Figure 1 provides the images of each component.

Age

Patient age is one of the first characteristics on the app input screen (Figure 1A). Although some participants recommended amending the age input field to include years and months rather than a calendar drop-down, most of their comments often focused on relevant treatment differences for young children or geriatric patients. These included the utility of mid-upper arm circumference (MUAC) measurement; availability of equipment for measuring blood pressure in children; and use of age-appropriate antibiotics, zinc, and vitamin A.

Danger Signs

The input screen also includes several specific danger signs. Temperature, entered in either degrees Celsius or degrees Fahrenheit, is recorded by choosing one of three radial buttons: <35, normal, and >37.9. A total of two *Yes* or *No* radial buttons were used for each of the following: fast breathing (defined as >40 breaths per minute [bpm]), unable to drink, and convulsions (Figure 1A). Participants found all of these data important for clinical judgment and diarrhea assessment. Suggestions for this screen included adding urine color, using a drop-down menu rather than *Yes* or *No* radial buttons for more precise measurement of ability to drink, and including the presence or absence of epigastric pain and comorbidities, such as diabetes or hypertension. A few others suggested that it would be important to have a means to record ongoing urine output, including the time last urine passed. Regarding the fast breathing field, some participants suggested that the respiratory rate cutoff >40 bpm was high. A participant's comments also suggested that they misunderstood *bpm* to reference a patient's pulse rate (ie, beats per minute).

The danger sign output screen provides algorithm-based recommendations using the input data. For example, when high fever is present, the app recommends "Check for sepsis or other causes." If the patient is vomiting and unable to drink, the app recommends "Space ORS sips or use IV fluids" (see the example in Figure 1B). Participants found the danger sign output important and relevant, with a participant commenting that sepsis cannot be properly diagnosed with the limited information used by the app.

Dehydration Assessment, Treatment Recommendations, and the Fluid Deficit Bar

The NIRUDAK app assesses dehydration according to what is entered in the dehydration assessment section of the input page (Figure 1A). Predictors included in this assessment are dependent on the selected NIRUDAK models (Multimedia Appendix 1 [23]) but generally included the following variables: eye level (normal or sunken), radial pulse (normal, decreased, or absent), vomiting in the last 24 hours (none, 1-5, 6-10, or >11 times), respiration depth (normal or deep), and skin pinch (rapid, slow, or very slow). A participant suggested adding the number of times stool has been passed to this section of the input page. Other responses to this screen included that it would take a little bit of time to choose from among the drop-down menu choices but that having the choices helped describe *a full, clear scenario* of the patient. In this discussion, another participant stated:

> *...[this app] is a good effort and [for] those who do not know [how to assess dehydration], they, by using the app, can do many things. [For] many people, [being in a] life threatening [situation] can be avoided. It is a great effort and a beautiful process; very good, I like the [app].* [government hospital nurse]

Using the information provided in the dehydration assessment section of the input screen, the output screen indicates whether the patient has *some*, *moderate*, or *severe* dehydration and whether danger signs are present. It also makes a recommendation about whether treatment in a hospital is needed, indicates the percent body weight lost, and provides a horizontal bar indicating the percentage of fluid deficit (Figure 1B). Participants discussed this output screen at length, many indicating that it was helpful and would be easy to use and/or

understand. A few participants commented that training would be needed to ensure that users would be familiar with the deficit bar and would understand the output and how to use it in treatment. Other recommendations included increasing the font size as well as color coding the information in red, yellow, and green to draw attention to patient risk level and treatment location. A participant also suggested clearly demarcating lines for the percentages on the fluid deficit bar. When asked to choose between an output screen with the fluid deficit bar and one without it, participants who stated a preference all chose the model with the fluid deficit bar. In 1 focus group, nurses discussed how the dehydration percentages for some, moderate, and severe dehydration are related to the WHO treatment guidelines for acute diarrhea [34]. Participants also discussed the role of weight in the calculation and the output screen, noting that if the patient's weight was based on an estimate rather than an actual measurement, the deficit bar data could be less useful or even incorrect. When asked if the dehydration assessment output would be helpful, a participant said:

> *Yes, definitely. Yes, obviously [it] will be helpful because I am getting it absolutely readymade. I do not have to think that much…This is excellent, isn't it? Excellent, nothing else, absolutely first class.* [government hospital physician]

Barriers to and Facilitators of App Use

We asked participants what they thought it would be like to use the NIRUDAK app in clinical practice, who should use the app tool, and what would support app use. Participants identified a variety of potential barriers to and facilitators of app use.

Barriers to use included the time it could take to train clinicians to use the app and the requirement for MUAC and systolic blood pressure (SBP) measurements. MUAC and SBP are used in calculating treatment recommendations in the full NIRUDAK model; however, participants noted that some clinical environments do not have MUAC measurement tape and blood pressure cuffs available:

> *...many times...digital blood pressure for children are not available. In that case, getting these two things [MUAC, SBP] accurately will be a bit difficult if I want to use it for mass population...I think the things [MUAC, BP] are good, but to use in mass population is a bit difficult.* [specialty hospital physician]

Similarly, it is not always possible to know if the patient has any medication allergies or another required field, particularly if they are nonresponsive.

Factors that ease or facilitate app use included clinicians' current experience with and use of other web-based tools, including local electronic health record systems, and familiarity with touch screens and other clinical apps:

> *As we are used to using the SHEBA app, in that case for us, it will not take so much time [to learn how to use app]. But, at the community level, [there may be patient] rush over there or [limited] manpower. In that case, for them, [app use] might face some problems.* [specialty hospital physician]

Recognizing that the app could be used in a variety of contexts, we asked what it would be like to use the app via telemedicine during a cholera outbreak and if community health workers (CHWs) could use it. There was a difference of opinion about whether the app could be incorporated into telemedicine. Some participants felt telemedicine use was possible, whereas others cautioned that important symptoms, such as sunken eyes, cannot be properly assessed via telemedicine. Many participants saw the app as useful for quick assessment and diagnosis, relatively easy to learn, and not requiring too much time to use, all of which would make it particularly useful during outbreaks. In contrast, some participants noted that time is further limited by the high numbers of patients requiring treatment during a cholera epidemic. In addition in that context, intravenous rehydration is usually started immediately upon arrival. As such, once rehydration has begun, there could be less need and time for assessment via the app, especially if MUAC and SBP measurements were required. The ability to calculate the recommended treatment without MUAC and blood pressure fields was considered a useful option for this context.

Opinion was divided on whether CHWs could use the app as a CDS tool. Although several participants, of which many were nurses, felt that CHWs would be able to use the app, others—often physicians—cautioned that clinical experience was essential and could be a limiting factor to nonphysician use of the app. Those who felt CHWs could use the app noted that it would make treatment decisions easy:

> *...to input information will be very easy. It may take 1 minute, or 2 minutes, or 3 minutes. Therefore, if health workers are trained, I think that they all can use this app...to make treatment decision.* [government hospital physician]

This could be a benefit in areas where physicians are less available. Caveats to CHW use included a concern that they would need to be trained to avoid mistreatment and to specifically observe the danger signs that indicate when hospital-based treatment is necessary.

Participants emphasized the relevance and importance of considering the end user's clinical expertise and judgment in two main ways: first, clinical experience is needed to support the use of the app through the accurate assessment of clinical signs, and second, the importance of avoiding overreliance on the app for clinical decision-making. Participants recognized that clinicians with varying levels of experience may interact differently with the app and that those with less experience may not recognize situations in which the app is less accurate or when there may be a degree of subjectivity (eg, assessment of clinical signs, such as sunken eyes). Similar comments were made by participants about the use of other formal clinical guidelines, which were felt to be primarily used by less experienced clinicians, whereas more experienced clinicians do not rely on guidelines as heavily: "We really do not treat people by [only using] guidelines in front of us, we [also] use our clinical judgement" [government hospital physician].

Considerations of Overtreatment and Undertreatment: Guidelines for the App's Clinical Recommendations

Physicians were shown the WHO's Integrated Management of Adolescent and Adult Illness (IMAI) guidelines for the classification of dehydration [21]. These use the presence of 2 or more clinical signs to classify dehydration as *no*, *some*, or *severe* and recommend appropriate treatment based on the dehydration classification [21]. Some participants, chiefly those from icddr,b, referenced the use of existing guidelines for the treatment of diarrhea, including the WHO guidelines and DHAKA method [22].

Physicians were also provided with data about the likelihood of correct treatment, overtreatment, and undertreatment of dehydration using the WHO's IMAI algorithm, and with 3 possible prediction cut points for classifying patients with severe dehydration in the NIRUDAK models (Figure 2). Each of these cut points corresponds to potential sensitivity and specificity thresholds for the NIRUDAK models. Sensitivity refers to the probability that the app will classify a patient as severely dehydrated when the patient is truly severely dehydrated (ie, true positive). Specificity refers to the probability that the app will classify a patient as not severely dehydrated when they are truly not severely dehydrated (ie, true negative) [25,35-37]. As presented in Figure 2, NIRUDAK option 1 illustrates a more sensitive model, selecting cutoffs in which of 100 patients, 57 (57%) would be correctly treated, 1 (1%) would be undertreated, and 42 (42%) would be potentially overtreated. This option avoids undertreating severe dehydration but may misclassify some patients who are not truly severely dehydrated. By contrast, NIRUDAK option 3 illustrates a more specific model, selecting cutoffs in which of 100 patients, 73 (73%) would be treated correctly, 3 (3%) would be undertreated, and 24 (24%) would be overtreated. This third option, compared with the other 2, avoids overtreating nonsevere dehydration but may misclassify some severely dehydrated patients as being nonsevere. Option 2 illustrates a model that falls in between options 1 and 3, neither highly sensitive nor highly specific. Facilitators asked physician participants to discuss their preferences and to weigh the risks and benefits of possibly undertreating severe cases and overtreating cases in which dehydration is not severe.

Figure 2. Overtreatment or undertreatment diagram presented to participants during focus group discussions. IMAI: Integrated Management of Adolescent and Adult Illness; NIRUDAK: Novel, Innovative Research for Understanding Dehydration in Adults and Kids; WHO: World Health Organization.

In general, participants indicated that it is important to avoid both undertreating and overtreating patients. In discussing the importance of not missing cases of severe dehydration, the physicians listed the subjective nature of some of the signs and symptoms used by the app and noted that these differed among very young, middle-aged, and older participants.

Physician preference was generally split between options 1 and 2. There were also no differences between physicians from the two settings regarding option choice; 3 doctors from icddr,b and 4 from government hospitals preferred option 1, whereas 2 physicians from each setting preferred option 2. In addition, 2 participants declined to make a selection or argued that although missing a patient with severe dehydration is problematic, because the patient may die, overtreatment can cause loss and damage as well: "over treatment is as perilous as [being] dehydrated" [specialty hospital physician].

Discussion

Principal Findings

Participants in the 8 focus groups were mostly enthusiastic about the NIRUDAK app, a novel CDS tool for diarrheal management

in low-resource settings. They highlighted the potential for time saving and the utility of the product during high-volume patient periods, such as during a cholera outbreak. Participants' opinions about key components and the barriers to and facilitators of app use were shared with the coinvestigators and the app design team and informed the next stage of the app development.

Factors that support the use of the NIRUDAK app include clinicians' familiarity with, and current use of, other mHealth tools, including a touch screen or haptic electronic medical record and other phone-based clinical apps. Notably, discussions about the anticipated utility of the app often occurred during FGDs with nurses, which could reflect a greater role that CDS tools may have for nurses than for physicians. This is an important consideration for the scale-up of the use of the app, given that a large number of patients with acute diarrhea and other common illnesses in LMICs are attended to by nurses or nonphysician health workers working in health centers rather than by physicians in hospital settings [38]. These findings suggest that targeting the use of the app toward nurses or nonphysician clinicians, especially those working in health facilities with lower resources, may allow the app to have the greatest impact.

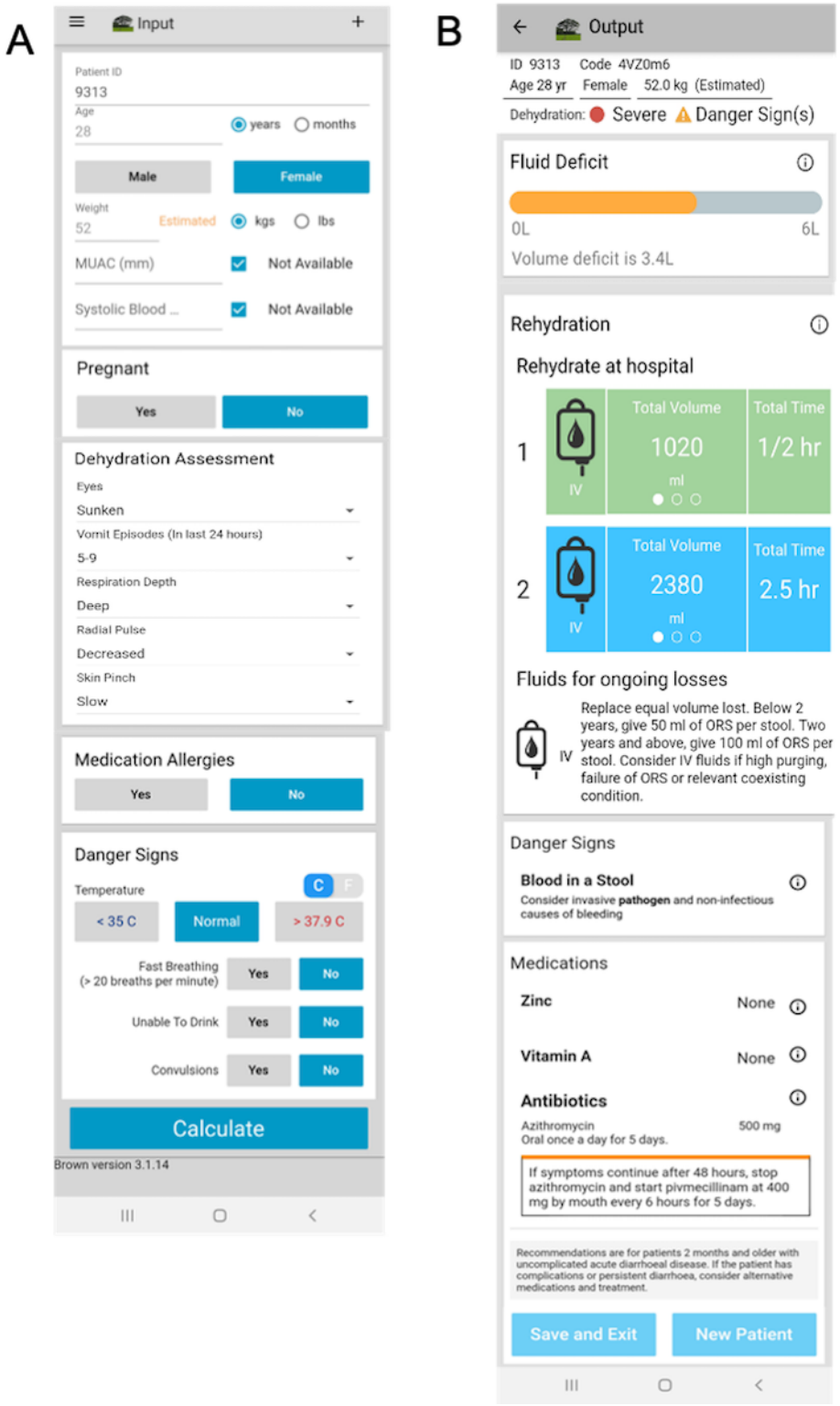
However, several physicians, cautioned against the use of the app by those with no or little formal clinical training, such as CHWs, versus those with formal clinical training, such as nurses and physicians. Such concerns were due to the possible misuse of the app or that CHWs may not be able to adapt the app's recommendations to unique cases. Further guidance and training in the use of the app and assessment of clinical signs may be important for the implementation of the app among CHWs and is in line with other recommendations that improving knowledge and skills is essential to improving the quality of care provided in LMICs [39].

The NIRUDAK app uses patient data to provide rehydration and other treatment recommendations based on whether the entered information indicates that the patient has no, some, or severe dehydration. The focus group questions about participants' use of existing treatment models, specifically the WHO's IMAI and the DHAKA method, were designed to help the researchers understand clinicians' opinions and treatment practices and to guide the models used in the NIRUDAK app itself. Of note, the overtreatment and undertreatment models shown in Figure 2 were provided as an aid during the FGDs so that the physicians could visually see how the various prediction cut points potentially used by the app would affect patient treatment. Our participants were concerned about the use of MUAC and SBP as metrics in calculating the rehydration recommendation. Although participants from the diarrheal specialty hospital indicated that these tools are sometimes used and may be available, participants from both the specialty hospital and the government hospitals cautioned that MUAC and SBP measurement will likely be unavailable in government hospitals or in more remote treatment clinics. Although the NIRUDAK app includes separate models based on resource availability, in response to this concern, the developers ensured the app automatically transitions to the simplified model for dehydration assessment when the user selects *Not Available* for either MUAC or SBP (Figure 3A). Such a feature was highly desirable by clinicians and will be appropriate for the app's use in contexts similar to Bangladesh at both specialty diarrhea and government hospitals. In general, CDS and mHealth tools in LMIC contexts must always consider resource limitations and allow for adaptation within the tool itself depending on resource availability [40].

The app developers made several other changes based on focus group participants' feedback (Figure 3). The age input field has been amended to include years and months for children aged <5 years. In addition, as 1 participant misunderstood *bpm* on the input page, this acronym was changed to its expanded form *breaths per minute* to prevent further confusion (Figure 3A). The development team was also concerned that the danger signs screen potentially required too much time and considered eliminating it. However, as the participants found it important and relevant, the danger signs section remained on the output screen (Figure 3B). After favorable feedback from participants, the fluid deficit bar, which was also under consideration for elimination, was not only retained but also redesigned in line with participants' reflections. The participants felt that it provided a necessary visual interpretation to help understand both the severity of a patient's dehydration and how much fluid they needed as part of their management (Figure 3B). In addition, it is also an improvement over the current WHO guidelines, which do not provide patient-specific guidance on how much fluid to give [21]. Finally, based on the feedback from the participants, option 2 was chosen as the default for predicting dehydration severity. However, because of the concerns expressed by the participants on overtreating or undertreating patients, an additional option was added to the settings menu of the NIRUDAK app that allowed clinicians to switch to the more sensitive option 1, which minimizes undertreatment, or the more specific option 3, which minimizes overtreatment, based on their practice settings and individual patient factors. The fact that there were no patterns of difference or preference between the 2 clinical settings further supported the decision to have both options available in the app.

Participant comments about the varying contexts of use and the varying needs of users indicate that this decision support tool for diarrhea treatment will be used differently depending on the clinician's role and the clinical context. For nurses and CHWs, the treatment recommendations provided through the app may be directive or proscriptive. For physicians and other advanced practitioners with significant diarrheal treatment experience, the app will support their clinical experience and judgment. Allowing clinicians to adjust the app settings, choosing between a more sensitive, more specific setting and the default setting both preserves the physician autonomy and allows for flexibility, making the app more generalizable to different clinical contexts, including when cholera is epidemic. In addition to the design choice allowing users to flexibly adapt the sensitivity and specificity of treatment recommendations, information tabs, which provide details on the model used, have been added to the app.

Figure 3. Screenshots of the NIRUDAK app's (A) input screen and (B) output screen after participants provided detailed feedback on the app. NIRUDAK: Novel, Innovative Research for Understanding Dehydration in Adults and Kids.



Limitations

These qualitative data about the clinical utility of the NIRUDAK app come from focus groups in which participants were shown still images of the app prototype. Consequently, feedback is limited to opinions about the appearance and content of the app; it is not based on actual use. Although conducting FGDs virtually using the Zoom platform allowed for the completion of data collection during the COVID-19 pandemic, poor internet connectivity prevented 19% (5/27) of the participants (4/5, 80%, were from government hospitals) from attending the focus groups, which may have influenced the findings. Meanings may have been lost in the translation process or interpretation of the data during the analytic process and could have introduced biases. However, the involvement of >1 researcher during the transcription, translation, and coding processes minimizes the

likelihood of misinterpreting research findings. In addition, this study was conducted in urban or semiurban hospital settings. Future work should focus on evaluating an mHealth app's clinical utility in rural or outpatient or ambulatory settings, as well as in other countries, and would be especially valuable.

Conclusions and Future Directions

The NIRUDAK app has been revised based upon formative qualitative data, which have contributed to the app's development and programming. The current iteration has been programmed with several features that were influenced by focus group participant feedback, including an option for clinicians to change between 2 different dehydration treatment models. The NIRUDAK app will be field-tested at icddr,b in 2022 to validate those models. Additional qualitative data will also be collected via individual interviews with nurses and physicians who field-test the app to further evaluate NIRUDAK's usability and understand the clinical users' experiences.

Acknowledgments

The authors would like to thank all the participants and study staff at the International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka Hospital, for their help and support. Funding was provided through grants from the National Institutes of Health, National Institute for Diabetes and Diarrheal and Kidney Diseases (principal investigator: ACL; grant 1R01DK116163-01A1). The funders had no role in the study design, data collection, or reporting processes. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the views of the National Institutes of Health or any governmental bodies or academic organizations.

Authors' Contributions

RKR, EJN, and ACL contributed to the study concept. RKR, ACL, SCG, MG, SN, and EJN contributed to the study design. SS, TH, SN, MG, RKR, NE, and SCG contributed to the data collection. SS, TH, and SN contributed to focus group discussion transcription and translations. SN, MG, and RL contributed to transcript editing. RKR, SCG, MG, and RL contributed to data analysis and the drafting of the manuscript. All authors edited and approved the final draft of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Predictors included in full and simplified Novel, Innovative Research for Understanding Dehydration in Adults and Kids models [23].
[DOCX File , 29 KB - humanfactors_v9i1e33325_app1.docx]

References

1. Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. 2019. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2021-07-21]
2. mHealth economics 2017 - current status and future trends in mobile health. Research 2 Guidance. 2017. URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2021-06-21]
3. mHealth economics 2017/2018 - connectivity in digital health. Research 2 Guidance. 2018. URL: https://research2guidance.com/product/connectivity-in-digital-health/ [accessed 2021-06-21]
4. Clinical decision support. HealthIT. URL: https://www.healthit.gov/topic/safety/clinical-decision-support [accessed 2021-06-21]
5. HIMSS mobile technology survey: mobile technology upgrades continue to impact healthcare engagement. Electronic Health Reporter. 2015. URL: https://electronichealthreporter.com/himss-mobile-technology-survey-mobile-technology-upgrades-continue-to-impact-healthcare-engagement/ [accessed 2021-06-21]
6. Pennic F. Infographic: top physician practice challenges and cost pressures. Health Information Technology Consultant. 2014. URL: https://hitconsultant.net/2014/03/10/infographic-top-physician-practice-challenges-and-cost-pressures/#.YPjaK24pDOQ [accessed 2021-06-21]
7. Patel R, Green W, Shahzad MW, Larkin C. Use of mobile clinical decision support software by junior doctors at a UK teaching hospital: identification and evaluation of barriers to engagement. JMIR Mhealth Uhealth 2015;3(3):e80 [FREE Full text] [doi: 10.2196/mhealth.4388] [Medline: 26272411]
8. Putzer GJ, Park Y. Are physicians likely to adopt emerging mobile technologies? Attitudes and innovation factors affecting smartphone use in the Southeastern United States. Perspect Health Inf Manag 2012;9(Spring):1b [FREE Full text] [Medline: 22737094]

9. Teferi GH, Tilahun BC, Guadie HA, Amare AT. Smartphone medical app use and associated factors among physicians at referral hospitals in Amhara region, North Ethiopia, in 2019: cross-sectional study. JMIR Mhealth Uhealth 2021;9(3):e19310 [FREE Full text] [doi: 10.2196/19310] [Medline: 33769303]
10. Peprah P, Abalo EM, Agyemang-Duah W, Budu HI, Appiah-Brempong E, Morgan AK, et al. Lessening barriers to healthcare in rural Ghana: providers and users' perspectives on the role of mHealth technology. A qualitative exploration. BMC Med Inform Decis Mak 2020;20(1):27 [FREE Full text] [doi: 10.1186/s12911-020-1040-4] [Medline: 32041608]
11. O'Reilly-Shah V, Easton G, Gillespie S. Assessing the global reach and value of a provider-facing healthcare app using large-scale analytics. BMJ Glob Health 2017;2(3):e000299 [FREE Full text] [doi: 10.1136/bmjgh-2017-000299] [Medline: 29082007]
12. Khan AI, Mack JA, Salimuzzaman M, Zion MI, Sujon H, Ball RL, et al. Electronic decision support and diarrhoeal disease guideline adherence (mHDM): a cluster randomised controlled trial. Lancet Digit Health 2020;2(5):e250-e258 [FREE Full text] [doi: 10.1016/S2589-7500(20)30062-5] [Medline: 33328057]
13. Portz JD, Ford KL, Elsbernd K, Knoepke CE, Flint K, Bekelman DB, et al. "i like the idea of it…but probably wouldn't use it" - health care provider perspectives on heart failure mHealth: qualitative study. JMIR Cardio 2020;4(1):e18101 [FREE Full text] [doi: 10.2196/18101] [Medline: 32885785]
14. Quan AM, Stiell I, Perry JJ, Paradis M, Brown E, Gignac J, et al. Mobile clinical decision tools among emergency department clinicians: web-based survey and analytic data for evaluation of the Ottawa rules app. JMIR Mhealth Uhealth 2020;8(1):e15503 [FREE Full text] [doi: 10.2196/15503] [Medline: 32012095]
15. Nair AA, Afroz S, Ahmed BU, Ahmed UU, Foo CC, Zaidan H, et al. Smartphone usage among doctors in the clinical setting in two culturally distinct countries: cross-sectional comparative study. JMIR Mhealth Uhealth 2021;9(5):e22599 [FREE Full text] [doi: 10.2196/22599] [Medline: 33970119]
16. Annual Report 2019 Solving public health problems through innovative scientific research. icddrb. 2019. URL: https://www.icddrb.org/dmdocuments/AR2019_17Jun2020.pdf [accessed 2022-03-21]
17. Ahmed S, Hasan MZ, Laokri S, Jannat Z, Ahmed MW, Dorin F, et al. Technical efficiency of public district hospitals in Bangladesh: a data envelopment analysis. Cost Eff Resour Alloc 2019;17:15 [FREE Full text] [doi: 10.1186/s12962-019-0183-6] [Medline: 31367193]
18. National Collaborating Centre for Women's and Children's Health (UK). Diarrhoea and vomitting caused by gastroenteritis: diagnosis, assessment and management in children younger than 5 years. London, UK: RCOG Press; 2009.
19. King CK, Glass R, Bresee JS, Duggan C, Centers for Disease Control and Prevention. Managing acute gastroenteritis among children: oral rehydration, maintenance, and nutritional therapy. MMWR Recomm Rep 2003;52(RR-16):1-16 [FREE Full text] [Medline: 14627948]
20. Acute diarrhea in adults and children: a global perspective. World Gastroenterology Organisation. 2012. URL: https://www.worldgastroenterology.org/UserFiles/file/guidelines/acute-diarrhea-english-2012.pdf [accessed 2021-04-01]
21. IMAI district clinician manual: hospital care adolescents and adults. World Health Organization. 2012. URL: https://apps.who.int/iris/rest/bitstreams/109693/retrieve [accessed 2021-06-21]
22. Levine AC, Glavis-Bloom J, Modi P, Nasrin S, Atika B, Rege S, et al. External validation of the DHAKA score and comparison with the current IMCI algorithm for the assessment of dehydration in children with diarrhoea: a prospective cohort study. Lancet Glob Health 2016;4(10):e744-e751 [FREE Full text] [doi: 10.1016/S2214-109X(16)30150-4] [Medline: 27567350]
23. Levine AC, Barry MA, Gainey M, Nasrin S, Qu K, Schmid CH, et al. Derivation of the first clinical diagnostic models for dehydration severity in patients over five years with acute diarrhea. PLoS Negl Trop Dis 2021;15(3):e0009266 [FREE Full text] [doi: 10.1371/journal.pntd.0009266] [Medline: 33690646]
24. Barry MA, Qu K, Gainey M, Schmid CH, Garbern SC, Nasrin S, et al. Derivation and internal validation of a score to predict dehydration severity in patients over 5 years with acute diarrhea. Am J Trop Med Hyg 2021;105(5):1368-1375. [doi: 10.4269/ajtmh.21-0143] [Medline: 34398821]
25. Fonseca BK, Holdgate A, Craig JC. Enteral vs intravenous rehydration therapy for children with gastroenteritis: a meta-analysis of randomized controlled trials. Arch Pediatr Adolesc Med 2004;158(5):483-490. [doi: 10.1001/archpedi.158.5.483] [Medline: 15123483]
26. Levine AC, Glavis-Bloom J, Modi P, Nasrin S, Rege S, Chu C, et al. Empirically derived dehydration scoring and decision tree models for children with diarrhea: assessment and internal validation in a prospective cohort study in Dhaka, Bangladesh. Glob Health Sci Pract 2015;3(3):405-418 [FREE Full text] [doi: 10.9745/GHSP-D-15-00097] [Medline: 26374802]
27. Haque F, Ball RL, Khatun S, Ahmed M, Kache S, Chisti MJ, et al. Evaluation of a smartphone decision-support tool for diarrheal disease management in a resource-limited setting. PLoS Negl Trop Dis 2017;11(1):e0005290 [FREE Full text] [doi: 10.1371/journal.pntd.0005290] [Medline: 28103233]
28. Garbern SC, Nelson EJ, Nasrin S, Keita AM, Brintz BJ, Gainey M, et al. External validation of a mobile clinical decision support system for diarrhea etiology prediction in children: a multicenter study in Bangladesh and Mali. Elife 2022;11:e72294 [FREE Full text] [doi: 10.7554/eLife.72294] [Medline: 35137684]

29. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. Int J Qual Methods 2019;18:160940691987459. [doi: 10.1177/1609406919874596]
30. Gray LM, Wong-Wylie G, Rempel GR, Cook K. Expanding qualitative research interviewing strategies: zoom video communications. Qual Rep 2020;25(5):1292-1301. [doi: 10.46743/2160-3715/2020.4212]
31. Guest G, MacQueen KM, Namey EE. Applied thematic analysis. Thousand Oaks, CA: Sage Publications; 2012.
32. NVivo 12. QSR International. 2018. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/support-services/nvivo-downloads [accessed 2022-03-14]
33. icddr,b pioneers use of technology to improve patient care. icddr,b. 2012. URL: https://www.icddrb.org/news-and-events/news?id=578&task=view [accessed 2021-08-24]
34. The treatment of diarrhoea: a manual for physicians and other senior health workers, 4th rev. World Health Organization. 2005. URL: https://apps.who.int/iris/handle/10665/43209 [accessed 2021-06-21]
35. Moons KG, van Es GA, Deckers JW, Habbema JD, Grobbee DE. Limitations of sensitivity, specificity, likelihood ratio, and Bayes' theorem in assessing diagnostic probabilities: a clinical example. Epidemiology 1997;8(1):12-17. [doi: 10.1097/00001648-199701000-00002] [Medline: 9116087]
36. Akobeng AK. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. Acta Paediatr 2007;96(3):338-341. [doi: 10.1111/j.1651-2227.2006.00180.x] [Medline: 17407452]
37. Ahs JW, Tao W, Löfgren J, Forsberg BC. Diarrheal diseases in low- and middle-income countries: incidence, prevention and management. Open Infect Dis J 2010;4:113-124. [doi: 10.2174/1874279301004010113]
38. Vasan A, Mabey DC, Chaudhri S, Brown Epstein HA, Lawn SD. Support and performance improvement for primary health care workers in low- and middle-income countries: a scoping review of intervention design and methods. Health Policy Plan 2017;32(3):437-452 [FREE Full text] [doi: 10.1093/heapol/czw144] [Medline: 27993961]
39. Bhutta ZA. Can technology help improve diarrhoea management? Lancet Digit Health 2020;2(5):e214-e215 [FREE Full text] [doi: 10.1016/S2589-7500(20)30081-9] [Medline: 33328052]
40. Biswas D, Hossin R, Rahman M, Bardosh KL, Watt MH, Zion MI, et al. An ethnographic exploration of diarrheal disease management in public hospitals in Bangladesh: from problems to solutions. Soc Sci Med 2020;260:113185 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113185] [Medline: 32712557]

Abbreviations

bpm: breaths per minute
CDS: clinical decision support
CHW: community health worker
DHAKA: Dehydration: Assessing Kids Accurately
FGD: focus group discussion
HIC: high-income country
icddr,b: International Centre for Diarrhoeal Disease Research, Bangladesh
IMAI: Integrated Management of Adolescent and Adult Illness
LMIC: low- and middle-income country
mHealth: mobile health
MUAC: mid-upper arm circumference
NIRUDAK: Novel, Innovative Research for Understanding Dehydration in Adults and Kids
SBP: systolic blood pressure
WHO: World Health Organization

Edited by A Kushniruk; submitted 10.09.21; peer-reviewed by K Taylor, B Chaudhry, D Carvalho; comments to author 30.11.21; revised version received 20.12.21; accepted 10.01.22; published 25.03.22.

Please cite as:
Rosen RK, Garbern SC, Gainey M, Lantini R, Nasrin S, Nelson EJ, Elshabassi N, Alam NH, Sultana S, Hasnin T, Qu K, Schmid CH, Levine AC
Designing a Novel Clinician Decision Support Tool for the Management of Acute Diarrhea in Bangladesh: Formative Qualitative Study
JMIR Hum Factors 2022;9(1):e33325
URL: https://humanfactors.jmir.org/2022/1/e33325
doi:10.2196/33325
PMID:35333190

Original Paper

Key Challenges and Opportunities for Cloud Technology in Health Care: Semistructured Interview Study

Kathrin Cresswell[1], PhD; Andrés Domínguez Hernández[2], PhD; Robin Williams[3], PhD; Aziz Sheikh[1], MD

[1]Usher Institute, The University of Edinburgh, Edinburgh, United Kingdom

[2]Department of Computer Science, University of Bristol, Bristol, United Kingdom

[3]Institute for the Study of Science, Technology and Innovation, The University of Edinburgh, Edinburgh, United Kingdom

Corresponding Author:
Kathrin Cresswell, PhD
Usher Institute
The University of Edinburgh
Teviot Place
Edinburgh, EH8 9AG
United Kingdom
Phone: 44 131 651 4151
Email: Kathrin.Cresswell@ed.ac.uk

Abstract

Background: The use of cloud computing (involving storage and processing of data on the internet) in health care has increasingly been highlighted as having great potential in facilitating data-driven innovations. Although some provider organizations are reaping the benefits of using cloud providers to store and process their data, others are lagging behind.

Objective: We aim to explore the existing challenges and barriers to the use of cloud computing in health care settings and investigate how perceived risks can be addressed.

Methods: We conducted a qualitative case study of cloud computing in health care settings, interviewing a range of individuals with perspectives on supply, implementation, adoption, and integration of cloud technology. Data were collected through a series of in-depth semistructured interviews exploring current applications, implementation approaches, challenges encountered, and visions for the future. The interviews were transcribed and thematically analyzed using NVivo 12 (QSR International). We coded the data based on a sociotechnical coding framework developed in related work.

Results: We interviewed 23 individuals between September 2020 and November 2020, including professionals working across major cloud providers, health care provider organizations, innovators, small and medium-sized software vendors, and academic institutions. The participants were united by a common vision of a cloud-enabled ecosystem of applications and by drivers surrounding data-driven innovation. The identified barriers to progress included the cost of data migration and skill gaps to implement cloud technologies within provider organizations, the cultural shift required to move to externally hosted services, a lack of user pull as many benefits were not visible to those providing frontline care, and a lack of interoperability standards and central regulations.

Conclusions: Implementations need to be viewed as a digitally enabled transformation of services, driven by skill development, organizational change management, and user engagement, to facilitate the implementation and exploitation of cloud-based infrastructures and to maximize returns on investment.

(JMIR Hum Factors 2022;9(1):e31246) doi:10.2196/31246

KEYWORDS

cloud technology; qualitative; adoption; implementation; digital health; data processing; health care; risk assessment; user engagement

Introduction

Background

There is now an international drive toward digitally enabled, data-driven transformation of health care services, with health systems seeking to optimize work processes; improve the quality, safety, and efficiency of care; and reduce costs [1,2]. Health care typically relies on a web of complex information infrastructures that lack integration and interoperability, which contributes to fragmented service provision [3]. Such infrastructures may range from systems allowing data analysis within individual organizations to advanced cloud-based systems facilitating cross-organizational data-driven analysis [4].

Although the origins of cloud technology can be traced back to the 1960s, the term *cloud computing* has only emerged in this millennium [5]. It essentially involves delegating storage and processing of data to third-party organizations accessed via the internet rather than hosting them on an organization's own computers. In doing so, cloud-based technologies can provide access to sophisticated large-scale technological infrastructures and advanced analytics services with the scope to rapidly scale up to meet peaks of demand [6]. Cloud product types differ in the degree of vendor and organizational control and can be public (shared across organizations), private (shared within organizations), or hybrid (a combination of both where on-premise infrastructure is combined with a public cloud). Hybrid clouds are increasingly popular as they not only allow access to public cloud infrastructure capacity but also maximize the use of on-premise solutions and therefore are a middle ground option for organizations with significant installed information technology (IT) capacity [7].

Textbox 1 summarizes the most common cloud products used in health care settings.

Textbox 1. Most common cloud products used in health care settings.

Common cloud products

- Software as a service, where a cloud provider hosts software services that user organizations can access on the web (eg, a cloud-based electronic health record such as Athenahealth)
- Platform as a service, where providers make development tools available to the user via the cloud (eg, Microsoft Azure)
- Infrastructure as a service, where the service provider supplies cloud-based infrastructure components to the client, such as storage, servers, and networks (eg, Virtustream Enterprise Cloud)

Although cloud computing has transformed many industries (eg, entertainment and financial services) [8], its use in health care remains limited. There are some exceptions of promising developments in advanced health care systems that are now reaping the benefits (Textbox 2) [9]. The advantages of the cloud have been particularly visible in the wake of the COVID-19 pandemic, which has called for rapid deployment and cross-organizational integration of services as well as large-scale real-time data analytics [10].

Textbox 2. Examples of advanced health care systems that have implemented cloud technology.

Examples of advanced health care systems that have implemented cloud technology

- The Shulan Health Management Group (China) implemented Amazon Web Services to host their "homegrown" system [11].
- The University of California, Los Angeles Health (United States) implemented Microsoft Azure for data processing and for integrating electronic health record data and data from other sources [12].
- The Mayo Clinic (United States) announced a strategic partnership with Google Cloud in 2019 [13].

However, despite some international governmental efforts to promote *cloud first* policies that foster the use of public cloud offerings in technology procurement [14,15], there are still significant points of friction in the adoption of cloud-based services. Some of these include concerns about security; fears of potential legal disputes between service providers and organizations; and issues surrounding vendor lock-in, privacy, ethics, and data ownership [16-20].

Objectives

In this study, we seek to understand how current opportunities in data-driven innovation facilitated by cloud computing could be positively harnessed in health care settings while minimizing perceived or actual risks.

Methods

Overview

We conducted a qualitative study between September 2020 and November 2020 using semistructured interviews sampling cloud providers, system implementers, software vendors, customers, and health informatics academics to gain an in-depth understanding of the evolving cloud ecosystem. It is important to keep in mind that data collection took place in the midst of the global COVID-19 pandemic and in the context of ongoing deliberations on the potential uses of cloud technology to address emerging urgent pandemic-related challenges. Discussions were strongly influenced by this topic.

Ethical Approval

We obtained ethical approval before the start of the study from the Usher Institute Research Ethics Group at the University of Edinburgh. Participants were provided with a consent form and an information sheet describing the study aims, procedures, and data management practices before participating in the study. They were given at least 48 hours to consider whether they agreed to participate and provided written informed consent. We informed the participants that they were free to withdraw at any time and that their responses would be anonymized during the analysis, removing names and places that could lead to identification of individuals.

Recruitment of Participants

We purposefully sampled stakeholders with perspectives on the topic of implementation, adoption, and optimization of cloud technology in health care settings [21]. Our aim was to gain a broad overview of different perspectives to understand the challenges and opportunities around cloud technology in health care settings and draw lessons that could inform future strategies for decision makers. In doing so, we specifically targeted individuals working across technology implementation, operations, design, research, and innovation within a range of organizations. We identified and recruited participants through our existing networks and communication channels as well as Google and LinkedIn searches using keywords related to the cloud and eHealth (eg, *digital health*, *digital transformation of health*, *cloud computing*, and *cloud first*). We complemented this strategy through snowball sampling by asking participants for recommendations of further interviewees. We aimed for variability in terms of geographical location (not including low- and middle-income countries as existing information infrastructures and challenges in these countries are likely to vary significantly), organizational function, area of expertise, and gender. Participants were selected based on their relationship with cloud technology in health care, both from the supply (cloud and software vendors) and demand (health care providers) sides. This included those who had experiences and opinions on the topic through experience of developing cloud solutions and cloud-enabled software, implementing and operating systems, or researching cloud technology.

Data Collection

ADH, a researcher with a background in science and technology studies and theoretical foundations surrounding information infrastructures, conducted all interviews via videoconference call software (Microsoft Teams). Interviews took the format of a *conversation with a purpose* where participants were encouraged to discuss issues important to them. ADH and KC (a social scientist with a background in sociotechnical theory) met periodically throughout the data collection process to discuss emerging findings and modify key lines of inquiry.

The interviews ranged in duration from 40 to 70 minutes. There were 20 one-to-one interviews and 1 group interview with 3 participants. Where participants asked for a group interview, we accommodated this request as it was more convenient for the participants and allowed us to gain insights into their complementary perspectives simultaneously. Although questions were tailored to individual roles and modified in line with emerging findings, we followed a topic guide exploring the state of cloud-enabled digital transformation in health care; views on barriers to realizing the potential benefits, risks, and areas of concern; and suggestions on how to address them (Textbox 3). During this process, the interviewer incorporated emerging themes across various interviewees and explored the tensions and differences in viewpoints in detail. We stopped collecting data when no new themes emerged during the concurrent analysis [22].

Textbox 3. Topic guide.

Topic guide

- Interviewee's background, current position, and description of the organization
- Overview of the cloud ecosystem, stakeholders, and existing offerings
- Implications of cloud adoption (cultural, organizational, operational, and financial adoption around digital transformation processes)
- Promising and concrete use cases of cloud technology in health care
- Challenges, risks, and hindrances for innovation in the cloud
- Distinctive challenges of health care compared with other industries and sectors
- Concerns about privacy, security, data ownership, and ethics
- State of affairs and challenges in terms of integration and interoperability between cloud platforms
- Role of the government
- Future outlook (5-10 years) of the cloud in health care

Data Analysis

The interviews were transcribed using an external professional service and subjected to thematic analysis [23]. ADH verified the interview transcripts by listening to the audio recordings and correcting any inaccuracies before analysis.

We used a mixture of deductive and inductive thematic coding [24]. We added the transcripts to an NVivo (QSR International) version 12 project and theme coded them using a sociotechnical coding framework developed by the research team [25]. This framework highlights how different technological and social dimensions interrelate and how different perspectives shape

aspects of the implementation and adoption of new technologies (Textbox 4). In addition, identified themes that did not fit the analytical framework were included in new categories.

Textbox 4. Dimensions used in the Technology, People, Organizations, and Macroenvironmental factors coding framework.

Dimensions used

- Technology (the technological properties of the system and the surrounding infrastructure)
- People (how various stakeholders use technology, including their expectations and experiences)
- Organizations (how organizations implement technology and how this shapes use)
- Macroenvironmental factors (how political and economic factors and markets shape technology development, use, implementation, and optimization)

ADH performed the first round of coding, periodically discussing emerging findings with KC. KC then re-examined the codes, resulting in minor changes to node titles and summarized the results in a narrative format. As part of our reflexive process, we identified how our previous experiences, assumptions, and preconceptions bore on the interpretation and coding of the data. In doing so, we discussed emerging findings within the research team to identify the relevance of themes within the Technology, People, Organizations, and Macroenvironmental factors (TPOM) framework as well as the need for new categories. We focused on examining converging and diverging perspectives, the interplay of technological and social dimensions, and the tensions and trade-offs emerging in the progress of cloud technology implementation, adoption, and optimization in health care settings.

Results

Overview

We interviewed 23 individuals (Table 1), including professionals working across major cloud providers, health care provider organizations, innovators, small and medium-sized software vendors, and academic institutions.

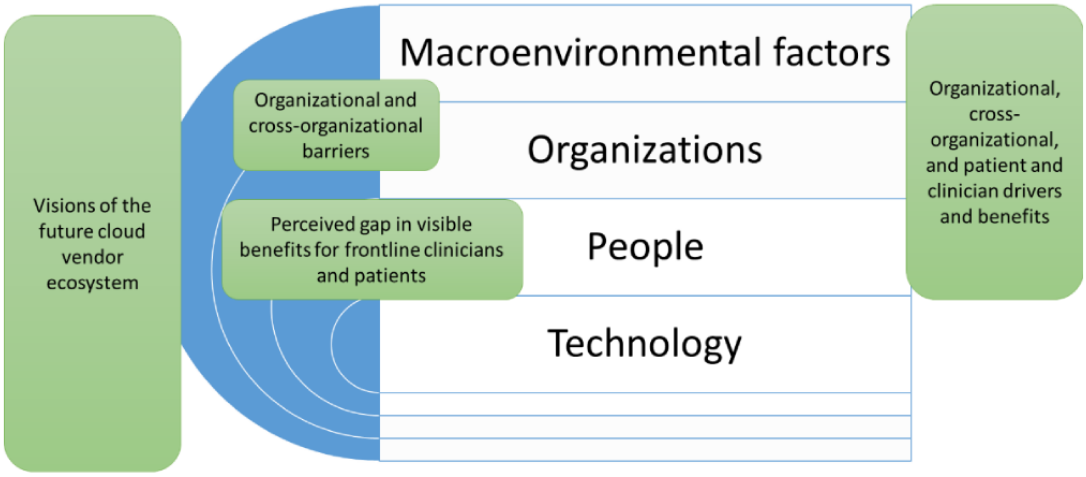
Table 1. Characteristics of the participants.

Participant number	Gender	Location	Occupation	Organization
1	Female	United States	Executive	Cloud vendor
2	Male	United Kingdom	Executive	Software vendor
3	Male	United Kingdom	Executive	Health care provider
4	Male	United Kingdom	Executive	Software vendor
5	Male	United Kingdom	Operations	Cloud vendor
6	Male	United Kingdom	Operations	Health care provider
7	Male	United Kingdom	Executive	Software vendor
8	Male	United Kingdom	Operations	Health care provider
9	Male	France	Operations	Cloud vendor
10	Male	United Kingdom	Operations	Cloud vendor
11	Male	United Kingdom	Academic	Health care provider
12	Female	United States	Operations	Cloud vendor
13	Female	United Kingdom	Operations	Cloud vendor
14	Female	Finland	Academic	Research
15	Male	United Kingdom	Operations	Software vendor
16	Female	United Kingdom	Executive	Research
17	Male	United Kingdom	Executive	Health care provider
18	Male	United Kingdom	Operations	Software vendor
19	Male	United Kingdom	Executive	Software vendor
20	Male	United Kingdom	Executive	Software vendor
21	Male	United Kingdom	Implementer	Health care provider
22	Female	United Kingdom	Operations	Cloud vendor
23	Male	United States	Executive	Cloud vendor

We produced 40 codes within the following four thematic areas: organizational context, social-human factors, technological factors, and wider macroenvironmental factors. The researchers then discussed the codebook and identified 4 salient challenges that were common across different interviewee backgrounds and affiliations. These were (1) drivers and perceived benefits associated with cloud technology in health care; (2) organizational and technological barriers limiting cost-effective use of cloud functionality; (3) infrastructural changes not immediately visible to frontline users, resulting in lack of clinical pull; and (4) visions of the future cloud vendor ecosystem.

Figure 1 illustrates how these emerging themes map onto the TPOM framework. As illustrated, the new emerging overarching categories related to cross-cutting issues spanning more than one TPOM dimension.

Figure 1. Overview of findings mapped onto the Technology, People, Organizations, and Macroenvironmental factors framework.



Drivers and Perceived Benefits of Cloud Technology in Health Care

The participants described various uses of cloud functionality, including scheduling software, videoconferencing, call center management, imaging analysis, and patient data analytics. On the basis of the most frequently discussed uses, we identified three salient categories: (1) organizational dimensions (eg, remote and collaborative working at scale, modeling algorithms surrounding predictive analytics, organizational analytics, and automation), (2) patient- and clinician-facing (eg, remote working, chatbots, and community outreach functionality), and (3) cross-organizational and regional (eg, data analytics surrounding particular disease areas for population health management and research).

We observed overall positive attitudes among the participants in relation to how cloud computing helped harness the value of data-driven innovation at scale. The adoption of cloud technology was perceived to be driven by existing issues faced by the sector, particularly concerning limited resources, access to and delivery of care, administrative workloads, and availability of critical services. Positive attitudes were particularly salient among system implementers, who saw immediate gains through secondary uses of data and tackling some of the most pressing challenges for health care posed by the COVID-19 pandemic. Here, cloud technology facilitated the deployment of solutions at speed without the need to purchase additional hardware:

> *[Cloud] enabled responsiveness...and throughout COVID, that's what that's been about. And it's removed one challenge [of] getting hold of hardware, getting it set up and all the rest. So, it's made us more responsive, it's made us quicker to adapt...the forcing function was COVID, and cloud's helped us have a faster response.* [Participant 2, male, software vendor, United Kingdom]

Other key benefits associated with cloud technology mentioned by the participants included cost-effective management and storage of data at scale combined with ready access to advanced computing capabilities and tools, such as machine learning (ML) and natural language processing:

> *And one of the greatest things about this now is machine learning and AI [artificial intelligence]...it hasn't been up until recently when [vendor] fully put a heavy effort over the last five or six years about democratizing access to these tools at scale, because you're not only interested in building one or two models, you're interested in building hundreds, thousands, tens of thousands of these models.* [Participant 23, male, cloud vendor, United States]

Organizational and Technological Barriers Limiting Cost-effective Use of Cloud Functionality

Although benefits associated with cloud technology were realized in many organizations that the participants worked with, they also discussed how these might not be representative of the wider health care landscape. Barriers manifested differently depending on the existing organizational and technological capabilities. Data migration and acquisition costs were mentioned by many interviewees from both the supply and demand sides. Cloud technology posed fewer barriers to organizations with few installed on-premise systems that sought to either implement new *pure* cloud-based solutions or rely on a *software as a service* business model. On the contrary, organizations with relatively mature digital infrastructures and

legacy systems faced hurdles to transition to cloud only solutions as they had to integrate existing systems and replace core infrastructures. Existing legacy systems were based on proprietary data structures and workflows, meaning that they could not simply be imported into the cloud. Instead, these organizations were more amenable to hybrid cloud solutions that relied on *infrastructure as a service* implementations:

> *The problem is, the cost of transition, if you're talking about your patient administration system or your electronic health record, which is often the core bit of software in your health care organization, if they want to switch that out, it is a huge job, which is massively expensive and massively risky to do.* [Participant 4, male, software vendor, United Kingdom]

In addition, implementers in particular raised the barriers associated with the need to change their cost structure with cloud technologies from capital up-front investments to a revenue model with recurring costs. This was perceived to be particularly problematic during the transitional period, when organizations were often running and paying for parallel systems:

> *In the short-term, you are inevitably paying more for the move towards cloud because you haven't necessarily got rid of all of that other infrastructure as you make that transition. So, you're now starting to pay for a revenue cost for your new cloud platform, but you've still got all of the cost of that other physical environment until you're able to decommission.* [Participant 16, female, research, United Kingdom]

Barriers not only related to cost but also to the organizational capabilities to adopt cloud solutions. Here, a lack of existing knowledge and skills in organizations to deploy and exploit cloud functionality was an important rate-limiting step. For instance, organizations frequently lacked implementation and migration skills:

> *In order to move things securely to the cloud either to implement brand new or to do a migration, you know, you need to have a certain degree of skill, knowledge, capability in order to do that...* [Participant 4, male, software vendor, United Kingdom]

Existing technical skills and capabilities also played an important role in maximizing the benefits of cloud functionality once it was implemented. Here, participants stated that many health care organizations lacked the knowledge and skills needed to work with advanced large-scale data analytics and therefore struggled to optimize the use of cloud infrastructures through artificial intelligence and ML:

> *There's a step that still needs to happen in the healthcare space, which is around just understanding what the analytics is.* [Participant 19, male, software vendor, United Kingdom]

Other barriers inhibiting uptake of cloud technologies in health care organizations included the changes in organizational culture required to transition to externally hosted systems and new modalities of accessing critical services. This was seen as particularly problematic for a risk-averse sector such as health care. For example, some participants mentioned that organizations that were skeptical about implementing cloud technology feared a loss of control if they migrated their IT systems to external service providers. In addition, there was apprehension about the reliability of the cloud and telecommunication infrastructure to deliver critical services, which manifested in the perceived need to fall back on on-premise IT services as contingency measures for critical services:

> *Traditionally, IT departments in [provider organizations], you have your server, you have your software on it, and they manage that. It makes them slightly uneasy if it's out there in a cloud and it's not something that they have control of.* [Participant 4, male, software vendor, United Kingdom]

Others stated that moving to cloud technology threatened established organizational hierarchies, particularly when sharing data across organizations. Health care settings were often not used to working across organizational boundaries. Cloud services challenged the traditional conception of organizations as autonomous entities and posed dilemmas in relation to information governance:

> *It's about [organizations] having to give up something to be part of a bigger collaboration.* [Participant 21, male, health care provider, United Kingdom]

Infrastructural Changes Not Immediately Visible to Frontline Users, Resulting in Lack of Clinical Pull

Although the organizational benefits of a wide range of cloud-based functions were visible and the case for organizational process and workload improvements could be made relatively easily by suppliers and system implementers, there was a perceived gap in visible benefits for frontline clinicians and patients. This presented a key barrier to the wider uptake of some cloud-based services as end users need to be on board for organizational changes to be implemented effectively:

> *An organizational imperative has to pass the challenge of the clinicians' view of what is important and vice versa. The clinicians' view of what is important has to pass the challenge of the gatekeepers in terms of organization, of funding, of development, service development, building development.* [Participant 18, male, software vendor, United Kingdom]

The underlying issue was the invisibility of digital infrastructures for those at the frontline, who mainly experienced benefits through the exploitation and optimization of these infrastructures once they were in place:

> *The people who are going to be using the technology, the people who are going to be using the insights from the analytics, the people who will be experiencing the change in process, they are almost don't really, it might sound harsh but...in the heat of the moment they almost don't really care about is it cloud enabled? What is the infrastructure? What's going*

> *on? Like, they just want the front end to work.* [Participant 19, male, software vendor, United Kingdom]

This lack of immediately visible benefits for end users combined with concerns surrounding privacy and security and the handling of sensitive data led to a lack of active user pull for cloud technologies in health care. It also presented challenges for suppliers as they had to satisfy a range of demands surrounding not only business processes but also clinical utility:

> *There's some particular challenges, how do you deal with the privacy aspects of the data and satisfy the concerns that data contributors and data custodians have, and then how do you accommodate for this enormous diversity within the user community in terms of how they use data and importantly how they get beyond very simple table analytics views of data into something that is more problematic, and how do you find a way for those outputs, those research outputs, to make their way back into clinical utility.* [Participant 7, male, software vendor, United Kingdom]

Despite these uneven perceptions, we also observed that during the COVID-19 crisis, clinical benefits and experiences of cloud technology became more common and thus immediately visible as remote consultations, remote working, data storage, and automation (eg, through chatbots) increasingly became a necessity:

> *Overnight we did see this huge uptick in the amount of telehealth, and that was only possible because of cloud there to support it.* [Participant 22, female, cloud vendor, United States]

Visions of the Future Cloud Vendor Ecosystem

Innovators, implementers, academics, and cloud vendors agreed on a vision characterized by a hybrid cloud-enabled ecosystem of applications where software suppliers rely on a combination of on-premise systems and cloud integration with a large cloud provider. For software suppliers, integrating with a cloud platform meant that they could quickly and cost-effectively scale up and scale down their products as required. This, in turn, was perceived to translate to lower risk and more efficient costing for health providers surrounding the trialing of new services:

> *For us, the main use cases are around working with a platform that allows us to quickly and cheaply get our product out into market...we don't need to invest huge amounts of time and people in developing things that are already out there...We can manage and maintain one environment, rather than having to think about how do you easily deploy and support, maintain, you know, 10, 20, 100 different customers, and the intricacies of deploying our app at every single customer site. We only have to think about one location.* [Participant 20, male, software vendor, United Kingdom]

However, the participants (in particular, implementers and software developers) also flagged the challenges and risks in terms of interoperability between different platform providers and integration between software vendors and cloud vendors. Innovators and system implementers voiced their expectations for interoperability standards and for cloud providers to open up application programming interfaces. However, opening up application programming interfaces and standardizing key functions was not always in line with legacy providers' commercial interests, which were typically based on retaining users within their platforms. Therefore, innovators in the software industry and implementers within health care organizations called for national regulations specifying interoperability standards to avoid vendor lock-in as this would allow for integration between systems and improve data portability. A lack of interoperability standards was viewed as inhibiting the development of a vibrant cloud ecosystem:

> *These regulatory bodies inside each of the governments would say the same thing, because that is the way to drive adoption of new technologies, forcing the new adoption, not rewriting everything, that's out of the question, but forcing for the benefit of all. I think this is how you're going to be having a government that is strong on that.* [Participant 9, male, cloud vendor, Europe]

Discussion

Principal Findings

Although the participants perceived clear drivers for the use of cloud technology in health care settings, particularly in relation to collaboration and workload efficiencies, barriers to progress included data migration costs and skill gaps within health care organizations to support implementation. This was exacerbated by the perceived cultural shift required to move to externally hosted services, challenging entrenched organizational ways of working and the need to reorganize existing cost structures. Frontline users, particularly those lacking technical expertise, were not directly concerned with the benefits associated with cloud-based infrastructures, which resulted in a lack of user pull in organizations seeking to change their technological infrastructures. However, the pressures of the COVID-19 pandemic and the stronger need for remote working arrangements made various critical cloud services visible. Central regulations and mandated interoperability standards were viewed as a key priority to foster innovation and reduce the risk of vendor lock-in.

Integration of Findings With the Current Literature

Our study confirms findings in other sectors that highlight that, despite the potential benefits, the move to cloud-based technologies in organizations necessitates cultural shifts from established ways of working and administering systems [26]. Therefore, it needs to be viewed as a complex sociotechnical transformation process, requiring not only technological but also socio-organizational changes to maximize the potential of cloud technologies [27]. Here, changes in organizational business models and technological infrastructures associated with cloud technology are likely to affect existing ways of working and organizational functioning as a whole [28]. Therefore, a key area of focus needs to be the effective

integration and embedding of new infrastructures with the *installed base* of existing technologies and socio-organizational structures and practices [4]. Barriers associated with data migration to cloud-based solutions are well documented in the literature [29], but our work also points to differences between digitally mature organizations with established installed technological systems (requiring more fundamental changes to the installed technological base) and those organizations that do not have established technological infrastructures, where data migration is likely to be less of an issue.

There is an asymmetry in the way system implementers, clinicians, and patients perceive and understand the benefits of the cloud, particularly when it comes to advanced functions such as ML and data-driven functionalities, which results in a lack of strong user pull [30]. User pull to implement cloud technologies within organizations is critical, especially in public service sectors [31]. Here, user attitudes and expectations toward technology can have a direct impact on adoption patterns [32]. A lack of perceived direct benefits as well as skepticism and concerns (most notably, perceived security, trust, and privacy issues) can result in negative attitudes toward a technology and lead to abandonment [33]. There are now growing calls for transparency and accountability of how personal data are used within cloud-based systems without compromising privacy and security [34]. Medical research is a key area where clinical data are considered immensely valuable but where handling of sensitive data is of utmost importance. This issue intersects not only with privacy and security but also with growing interest across industry and academia on trustworthy, fair, and ethical use of big data and algorithmic technologies [20,35,36]. Therefore, it is critical for organizations promoting the use of cloud technology to place emphasis on active engagement with users and rigorously engage with debates about privacy, ethics, and security taking place in academic and public forums [37].

The move to cloud technologies in health care presents a disruptive innovation for the market [30,38], which inevitably results in tensions and trade-offs between conflicting agendas and interests. In this study, we observed that points of friction related to the integration of different building blocks and interoperability between competing platforms. These challenges resonate with previous studies in information systems, which highlight ongoing tensions between requirements for standardization and the flexible and cost-effective operation of systems [39-41].

Implications for Policy, Practice, and Research

Among the key challenges voiced by our informants were the lack of installed capacity and technical skills, the cost of migration, and the need for investment restructuring. As a result, organizations that still rely on on-premise IT infrastructure and software see hybrid cloud solutions as a way forward. There is now a need to support the development of such hybrid structures and map potential integration and migration pathways to help implementing organizations envisage new information infrastructure constellations. This needs to be supported by active efforts to address the existing skill gap in cloud computing and digital transformation expertise in the health sector [42]. This will also help ensure that advanced cloud functions such as ML are effectively exploited.

Strategic decision makers need to recognize the need to view the implementation of cloud-based systems as a major digital transformation of services to promote cloud first policy in health care settings. Therefore, implementations need to be supported not only by technological capability but also by change management expertise and continuous stakeholder engagement.

Our work highlights divergent views and expectations among various stakeholder groups in relation to interoperability. These are highly contingent upon political-economic contexts as interoperability standards are not always centrally mandated across countries. Innovators and system implementers in particular raised the need to regulate the emerging cloud ecosystem through the development of interoperability standards. Adding to the risk of developing solutions for a particular vendor is poor integration between competing platforms. A clear policy recommendation to address this challenge is the central mandate for interoperability standards, with the United States being a case for reference, but these need to be flexible to respond to emerging needs and other disruptive innovations that are likely to emerge. Of central importance will be the need for trustworthy entities and tools for responsible use of sensitive data, developing mechanisms for ensuring ethical and transparent use for medical research without compromising patients' privacy and integrity.

Strengths and Limitations

We gained insights into the opportunities and challenges in the emerging area of cloud technology implementation in health care settings by consulting a range of perspectives. We deliberately sampled implementers, customers, academics, and vendors to explore experiences and insights from a range of settings. However, this may have been at the expense of breadth. For example, consumer and customer perspectives were underrepresented in our sample, and we did not consult the range of immediate frontline users of technologies or legal and privacy experts. Our sample also consisted mainly of cloud enthusiasts. Nevertheless, our study points to various user-facing issues such as adoption, use, concerns, and invisibility of functions, which we assessed indirectly through respondents working in close contact with users. Further empirical work with clinicians, lawyers, and privacy experts arises as a pertinent avenue of research.

Our themes provide a helpful guide for conducting future in-depth work as we have illustrated an overview of tensions. In addition, we would also have liked a broader representation of international settings (as 18/23, 78% of participants in our sample were based in the United Kingdom). Our current sample consisted mainly of participants from North America and Europe (France, Finland, and the United Kingdom). Future work should build on our findings seeking to explore how different geographies, including low- and middle-income countries, have approached the area and how challenges vary across different core infrastructures, levels of digital maturity, and health system organization.

Conclusions

Although cloud technologies promise to deliver a range of technical capabilities, they are unevenly applied across health care settings depending on organizational contexts and existing infrastructures. In the wake of the pandemic, cloud technologies have become vital to support everyday collaboration for clinicians, remote health delivery, and other operational functions, which has considerably driven the adoption of the cloud. Going forward, cloud implementation needs to be viewed as disruptive organizational change initiatives facilitated by national initiatives to promote interoperability for a vibrant cloud ecosystem. Areas that may lend themselves to such work may include patient-facing technologies, where cloud providers are already established, and health and social care integration, where limited existing health information infrastructures may reduce barriers associated with integration or migration. This will also need to involve engaging in public discourse about cost, risk, and trust (or lack thereof) in cloud platforms regarding the handling of sensitive data, privacy, security, and ethics.

Acknowledgments

The authors gratefully acknowledge all the participants' time and input. They also thank Scott Watson for his thoughtful comments on the paper and Rachel Dunscombe and John Halamka for helping with the recruitment of participants. This work was funded by a Scottish Government Chief Scientist Research Grant. The views expressed are those of the authors.

Conflicts of Interest

None declared.

References

1. Bates DW, Gawande AA. Improving safety with information technology. N Engl J Med 2003 Jun 19;348(25):2526-2534. [doi: 10.1056/nejmsa020847]
2. Healthcare big data and the promise of value-based care. NEJM Catalyst. 2018. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0290 [accessed 2021-12-14]
3. Olaronke I, Soriyan A, Gambo I, Olaleke J. Interoperability in healthcare: benefits, challenges and resolutions. Int J Innov Applied Stud 2013;3(1):262-270 [FREE Full text]
4. Aanestad M, Grisot M, Hanseth O, Vassilakopoulou P. Information Infrastructures Within European Health Care: Working With the Installed Base. Cham: Springer Nature; 2017. [doi: 10.1007/978-3-319-51020-0_3]
5. The history of cloud computing. Server Mania. URL: https://blog.servermania.com/the-history-of-cloud-computing/ [accessed 2021-12-14]
6. Almutiry O, Wills G, Alwabel A, Crowder R, WaIters R. Toward a framework for data quality in cloud-based health information system. In: Proceedings of the International Conference on Information Society (i-Society 2013). 2013 Presented at: International Conference on Information Society (i-Society 2013); Jun 24-26, 2013; Toronto, ON, Canada URL: https://ieeexplore.ieee.org/document/6636362
7. Mazhelis O, Tyrväinen P. Economic aspects of hybrid cloud infrastructure: user organization perspective. Inf Syst Front 2011 Sep 14;14(4):845-869. [doi: 10.1007/s10796-011-9326-9]
8. Schwertner K. Digital transformation of business. Trakia J Sci 2017;15(Suppl.1):388-393. [doi: 10.15547/tjs.2017.s.01.065]
9. Mizani M. Cloud-based computing. In: Key Advances in Clinical Informatics. Cambridge: Academic Press; 2017.
10. Cresswell K, Williams R, Sheikh A. Using cloud technology in health care during the COVID-19 pandemic. The Lancet Digital Health 2021 Jan;3(1):e4-e5. [doi: 10.1016/s2589-7500(20)30291-0]
11. Shulan Health. Amazon Web Services. URL: https://www.amazonaws.cn/en/customer-stories/health/shulan/ [accessed 2021-12-22]
12. UCLA health using Microsoft azure to improve EHR data integration. EHR Intelligence. URL: https://ehrintelligence.com/news/ucla-health-using-microsoft-azure-to-improve-ehr-data-integration [accessed 2021-12-14]
13. Mayo Clinic selects Google as strategic partner for health care innovation, cloud computing. Mayo Clinic. URL: https://newsnetwork.mayoclinic.org/discussion/mayo-clinic-selects-google-as-strategic-partner-for-health-care-innovation-cloud-computing/ [accessed 2021-12-14]
14. Cloud First is here to stay. GOV.UK. URL: https://technology.blog.gov.uk/2019/10/31/cloud-first-is-here-to-stay/ [accessed 2021-12-14]
15. From cloud first to cloud smart. Federal Cloud Computing Strategy. URL: https://cloud.cio.gov/ [accessed 2021-12-14]
16. Gray A. Conflict of laws and the cloud. Comput Law Security Rev 2013 Feb;29(1):58-65. [doi: 10.1016/j.clsr.2012.11.004]
17. Schukat M, McCaldin D, Wang K, Schreier G, Lovell N, Marschollek M, et al. Unintended consequences of wearable sensor use in healthcare. Contribution of the IMIA wearable sensors in healthcare WG. Yearb Med Inform 2016 Nov 10(1):73-86 [FREE Full text] [doi: 10.15265/IY-2016-025] [Medline: 27830234]
18. Sajid A, Abbas H. Data privacy in cloud-assisted healthcare systems: state of the art and future challenges. J Med Syst 2016 Jun;40(6):155. [doi: 10.1007/s10916-016-0509-2] [Medline: 27155893]

19. Kuo AM. Opportunities and challenges of cloud computing to improve health care services. J Med Internet Res 2011 Sep 21;13(3):e67 [FREE Full text] [doi: 10.2196/jmir.1867] [Medline: 21937354]
20. Duquenoy P, Mekawie N, Springett M. Patients, trust and ethics in information privacy in eHealth. In: eHealth: Legal, Ethical and Governance Challenges. Germany: Springer; 2013.
21. Coyne I. Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? J Adv Nurs 1997 Sep;26(3):623-630. [doi: 10.1046/j.1365-2648.1997.t01-25-00999.x] [Medline: 9378886]
22. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
23. Boyatzis R. Transforming Qualitative Information: Thematic Analysis and Code Development. Thousand Oaks, California: SAGE Publications; 1998.
24. Joffe H, Yardley L. Content and thematic analysis. Res Methods Clin Health Psychol 2004;56:68. [doi: 10.4135/9781849209793.n4]
25. Cresswell K, Williams R, Sheikh A. Developing and applying a formative evaluation framework for health information technology implementations: qualitative investigation. J Med Internet Res 2020 Jun 10;22(6):e15068 [FREE Full text] [doi: 10.2196/15068] [Medline: 32519968]
26. Sultan N, van de Bunt-Kokhuis S. Organisational culture and cloud computing: coping with a disruptive innovation. Technol Analysis Strat Manag 2012 Feb;24(2):167-179. [doi: 10.1080/09537325.2012.647644]
27. Liu Y, Dong S, Wei J, Tong Y. Assessing cloud computing value in firms through socio-technical determinants. Inf Manag 2020 Dec;57(8):103369. [doi: 10.1016/j.im.2020.103369]
28. Sørheller VU, Høvik EJ, Hustad E, Vassilakopoulou P. Implementing cloud ERP solutions: a review of sociotechnical concerns. Procedia Comput Sci 2018;138:470-477. [doi: 10.1016/j.procs.2018.10.065]
29. Phaphoom N, Wang X, Samuel S, Helmer S, Abrahamsson P. A survey study on major technical barriers affecting the decision to adopt cloud services. J Syst Softw 2015 May;103:167-181. [doi: 10.1016/j.jss.2015.02.002]
30. Grisot M, Hanseth O, Thorseng A. Innovation of, in, on infrastructures: articulating the role of architecture in information infrastructure evolution. J Assoc Inf Syst 2014 Apr;15(4):197-219. [doi: 10.17705/1jais.00357]
31. Shin D. User centric cloud service model in public sectors: policy implications of cloud services. Government Inf Q 2013 Apr;30(2):194-203. [doi: 10.1016/j.giq.2012.06.012]
32. Wu J, Shen W, Lin L, Greenes RA, Bates DW. Testing the technology acceptance model for evaluating healthcare professionals' intention to use an adverse event reporting system. Int J Qual Health Care 2008 Apr;20(2):123-129. [doi: 10.1093/intqhc/mzm074] [Medline: 18222963]
33. Laumer S, Eckhardt A. Why do people reject technologies: a review of user resistance theories. In: Information Systems Theory. New York: Springer; 2012.
34. Securing the E-Health cloud. ACM. URL: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=7C1DC7517674B9D7CDD09D023304FCA3?doi=10.1.1.717.462&rep=rep1&type=pdf [accessed 2021-12-14]
35. McCradden MD, Joshi S, Mazwi M, Anderson JA. Ethical limitations of algorithmic fairness solutions in health care machine learning. Lancet Digit Health 2020 May;2(5):e221-e223. [doi: 10.1016/s2589-7500(20)30065-0]
36. Ahmad M, Patel A, Eckert C, Kumar V, Teredesai A. Fairness in machine learning for healthcare. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2020 Presented at: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; Jul 6 - 10, 2020; CA USA. [doi: 10.1145/3394486.3406461]
37. Cresswell K, Morrison Z, Crowe S, Robertson A, Sheikh A. Anything but engaged: user involvement in the context of a national electronic health record implementation. Inform Prim Care 2011;19(4):191-206 [FREE Full text] [doi: 10.14236/jhi.v19i4.814] [Medline: 22828574]
38. Cloud computing as a disruptive technology. Cloud Academy. URL: https://cloudacademy.com/blog/disruptive-technology-and-cloud-computing/ [accessed 2021-12-14]
39. Bowker G, Baker K, Millerand F, Ribes D. Toward information infrastructure studies: ways of knowing in a networked environment. In: International Handbook of Internet Research. Dordrecht: Springer; 2009.
40. Cennamo C, Santaló J. Generativity tension and value creation in platform ecosystems. Org Sci 2019 May;30(3):617-641. [doi: 10.1287/orsc.2018.1270]
41. Understanding infrastructure: dynamics, tensions, and design. University of Michigan Library. URL: https://deepblue.lib.umich.edu/handle/2027.42/49353 [accessed 2021-12-14]
42. Foster D, White L, Adams J, Erdil D, Hyman H, Kurkovsky S, et al. Cloud computing: developing contemporary computer science curriculum for a cloud-first future. In: Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education. 2018 Presented at: Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education; Jul 2 - 4, 2018; Larnaca Cyprus. [doi: 10.1145/3293881.3295781]

Abbreviations

IT: information technology
ML: machine learning
TPOM: Technology, People, Organizations, and Macroenvironmental factors

Edited by A Kushniruk; submitted 15.06.21; peer-reviewed by J Walsh, I Wilson; comments to author 11.09.21; revised version received 14.09.21; accepted 02.10.21; published 06.01.22.

Please cite as:
Cresswell K, Domínguez Hernández A, Williams R, Sheikh A
Key Challenges and Opportunities for Cloud Technology in Health Care: Semistructured Interview Study
JMIR Hum Factors 2022;9(1):e31246
URL: https://humanfactors.jmir.org/2022/1/e31246
doi:10.2196/31246
PMID:34989688

Original Paper

Implementation of E-prescription for Multidose Dispensed Drugs: Qualitative Study of General Practitioners' Experiences

Monika Knudsen Gullslett[1], PhD; Trine Strand Bergmo[1], PhD

Norwegian Centre for E-Health Research, University Hospital of North Norway, Tromsø, Norway

Corresponding Author:
Monika Knudsen Gullslett, PhD
Norwegian Centre for E-Health Research
University Hospital of North Norway
Postboks 35
Tromsø, 9038
Norway
Phone: 47 90784208
Email: monika.knudsen.gullslett@ehealthresearch.no

Abstract

Background: Increased use of pharmaceuticals challenges both capacity and safety related to medication management for patients and changes in how general practitioners (GPs) and other health personnel interact with and follow up with patients. E-prescribing of multidose drug dispensing (eMDD) is 1 of the national measures being tested in Norway.

Objective: The objective of this study is to explore GPs' experiences with the challenges and benefits of implementing eMDD in Norway.

Methods: Qualitative in-depth and group interviews were conducted with a total of 25 GPs between 2018 and 2020. Transcribed files were saved in NVivo to conduct a step-by-step content analysis. NVivo is a software tool for organizing, managing, and analyzing qualitative data.

Results: The study revealed that eMDD offers many benefits. At the same time, there are several challenges related to information, training, and initiation, as well as to the responsibility for the medication, interactions, and the risk of incorrect medication. An important activity in the start-up phase was an information meeting with pharmacies and technology suppliers, as well as exchanging information and instructions with pharmacies on how to get started. Four analytic themes emerged through the extraction of data: (1) start-up with eMDD ("Be patient"); (2) the need for training; (3) interaction, safety, and efficiency; and (4) the working day with eMDD.

Conclusions: There is a variation in different GPs' needs regarding training and information, and considerable variation in competence and motivation related to the use of digital tools. There are also different degrees of understanding the everyday work of the other actors in the medication chain. In particular, the harmonization of medication lists related to the use of time, expenditures, and challenges with technological solutions in the introduction phase was emphasized as a challenge. Overall, GPs who have started using the system report great benefits; these are largely related to an increased overview of patients' total medication lists, less time spent on prescribing prescriptions, and increased collaboration with pharmacies and nurses, both in service from providers in homes and in nursing homes.

(JMIR Hum Factors 2022;9(1):e27431) doi:10.2196/27431

KEYWORDS

e-prescribing of multidose drug dispensing (eMDD); pharmacy; start-up; general practitioner (GP); Norway; digital health; digital tools; e-prescriptions; physicians; qualitative study

Introduction

Background

Digitalization and the use of electronic systems to manage medication are salient elements in developing future health care services that have a current political, clinical, and research focus [1,2]. With a population characterized by an increasing proportion of fragile and older people and people in need of health care, the use of different types of medication is also increasing [3-5]. Changes in population structures have also resulted in a need to change how general practitioners (GPs)

and other health personnel perform their tasks and how patients are medicated [6]. This also applies to the handling of medications for people who have several diagnoses (multimorbidity) and who are thus dependent on several types of medication (polypharmacy) [7-9]. In Norway, among other countries, a recent change has been the introduction of e-prescribing of multidose drug dispensing (eMDD) [10-12]. The goal of e-prescription technology is to contribute to a more conscious and safer use of multiple medications by a single individual [13,14]. It is expected that technology, such as the use of eMDD and other solutions, will reduce duplication of medications, contribute to correct dosages, and reduce confusion among providers and patients [12,15].

What Is MDD, and What Is eMDD?

MDD was introduced in the early 2000s to reduce errors and streamline the distribution of medications in municipal health services [6,8]. The main goal of introducing MDD was to reduce incorrect dispensing, save time for health providers, and reduce the disposal of medications [11,16,17]. MDD is intended to replace pill dispensers and is a mechanical system used by central pharmacies to package medications in small unit-of-use bags, with 1 bag for each dosing time. Users receive a strip with many small bags marked with the patient ID, medication information, and time of ingestion. MDD is packaged and delivered every fortnight. Today, more than 90,000 people use MDD in Norway [18]. Of these, 68,400 (76%) receive service from professional caregivers, 18,900 (21%) live in nursing homes, and just under 3600 (4%) receive multidose drugs by private agreement with a pharmacy [1]. GPs prescribe multidose drugs by listing the patient's medications on a prescription card. This medication list is then printed out and mailed or faxed to the pharmacy. Once a GP signs this medication list, it is valid as a prescription for 1 year.

Errors in the e-prescribing of medication as incorrect medication may be a serious problem [8,19]. Several issues, such as training staff, designing routines, and focusing on the environmental aspects of the practice, are important for avoiding such errors [19]. In Norway, safe digital routines are increasingly being designed and implemented to prevent incorrect medications, improve patient quality of life and safety, and contribute to a more efficient workday for practitioners, including GPs [7]. As part of the digitalization of health services, Norwegian health authorities have begun testing eMDD within the e-prescription solution [20]. In this paper, we focus on the challenges and benefits GPs experience when implementing eMDD.

Today, over 90% of prescriptions are sent to the pharmacy as e-prescriptions [21]. Among the medications still prescribed on paper are multidose drugs; in Norway, the goal is to transfer these to the e-prescription system. Toward this end, the Norwegian health authorities have begun testing eMDD within the e-prescribing system. eMDD means that an electronic medication list and e-prescriptions replace the paper list and fax. The GP sends a list of the patient's regular medications to the prescription database, together with e-prescriptions, as the medication list has no prescription function [22]. The prescription database is a central database where prescribing information is shared between all health personnel with prescribing privileges and all pharmacies. In the eMDD system used by the GPs in our study, an electronic MDD message that includes the medication lists was added to the e-prescriptions. The medication list for MDD is called the "medications in use" list (or just the patient's medication list). The patient's medication list shows the patient's regular medications, medications as needed, dietary supplements, and any critical information related to the medications, as well as recently discontinued medications. To be able to submit an MDD notification to the prescription database, the doctor must first register as an MDD-responsible doctor. The medication lists and e-prescriptions are developed in the GP's prescribing module in the medical record system and sent to the prescription database. The pharmacy can access the information and transfer it to the packing machine for dispensing. Prescriptions are displayed in the prescription database 1 month after they have expired. In the e-prescription system, the responsible doctor and the pharmacy electronically communicate regarding any necessary clarifications. The home care staff receives an e-message notification when GPs make changes to patients' medication lists [19]. The target group for eMDD in Norway is currently patients in municipal nursing and care services. There is some variation in the user groups related to age and, among other things, the use of medications.

Going from paper to electronic medication lists also makes medication information available in the prescription database to medical doctors in emergency rooms and hospitals, making medication information more available during care transitions. Moreover, the GP receives a notification from the pharmacy if there are changes in the medication treatment that have not been initiated by the GP [1].

Studies have highlighted a significant decrease in the number of discrepancies between the medication lists at GPs and pharmacies when eMDD is compared with MDD prescribed using paper and fax [22]. In addition, research related to digitalizing prescriptions has found that health professionals experience the solution as a quality improvement [23]. Results from another study focused on the potential to streamline workflow for health care providers and minimize interruptions from, among other things, the use of phone and fax communications. This study also emphasized that technical standards and system design changes, and more targeted training, may be needed to address barriers to e-prescriptions [24].

Context for This Study

The goal of the digital solution being implemented is to provide a safer system based on electronic routines and updated medication lists. As stated, health authorities expect eMDD to reduce incorrect use of drugs, but research in the field is sparse, and studies from other countries cannot be transferred directly to the Norwegian context due to other systems and routines for eMDD [17]; this study also found that health professionals experience the solution as a quality improvement [17]. Discrepancies in the medicine list are also a challenge in transition between institutions, or institutions and homes, in the processes of admission and discharge from the hospital, and

electronic tools may be helpful to avoid or minimize medication discrepancies [25].

eMDD was piloted in 24 GP clinics/offices in the southern part of Norway between 2018 and 2020. As part of this pilot, we undertook a study investigating how the GPs experienced the transition from prescribing MDD using paper and fax to eMDD. This paper presents findings from interviews conducted with GPs in different parts of the country based on the research question, *What are GPs' experiences with their challenges and benefits regarding the introduction, implementation, and use of eMDD in their practices?*

Methods

Research Design

A qualitative, explorative study using in-depth interviews was conducted in GP practice. The study's methodological approach was based in the social sciences, using an abductive strategy that aimed to uncover—and then interpret—knowledge about the social actors in question [26]. Different research strategies are summarized by Peirce [27]: "Deduction proves that something must be; induction shows that something actually is operative; abduction merely suggests that something may be." The abductive strategy works well with the hermeneutic-phenomenological approach we used in the analysis; moreover, the choice of research strategy was integrated into the study's objectives and the research questions under investigation. In this study, the choice of a hermeneutic-phenomenological perspective means that the researchers tried to achieve an in-depth understanding of the study participants' life-world experiences with the topic of the study and to uncover and interpret knowledge about GPs' experiences with implementing eMDD [28,29]. Even if the data gathering and analysis are done with a reflexive and open-minded view, the researcher's hermeneutic position will affect the results based on the theoretical approach and their preconceptions [30,31].

Selection, Sample, and Interviews

The findings in this paper build on this knowledge and focus on the experiences GPs have related to the implementation and use of eMDD. Experiences from the pilot eMDD and whether the system meets expectations were investigated. The Results section emphasizes the GPs' work situation and patient safety—that the patient receives the right medicine at the right time—and the analysis places the eMDD experiences in light of the complexity it is part of and also ensures patient safety and the correct use of medicines for patients who are prescribed medicines through eMDD. All GPs who implemented eMDD between 2018 and 2019 were invited to participate in the study. A total of 24 GP clinics/offices were involved in testing eMDD during this period. We received contact information for these offices from the Directorate for e-Health in Norway. A total of 26 GPs from 10 doctors' offices agreed to participate, of which 3 (11%) agreed to participate in a follow-up interview. A qualitative, explorative study using in-depth interviews was conducted by both authors to investigate how GPs experienced the introduction of eMDD. A third researcher (EJ) conducted interviews together with researcher TSB.

Two focus groups were conducted, with 8 doctors in each interview and 9 individual interviews, 4 of which were telephone interviews. In addition, 1 GP described experiences with eMDD in 2 e-mails. The GPs had between 5 and 20 patients who used MDD. Some GPs were salaried, while others worked on a contractual basis. Some were interviewed after 2-4 months of use and others after 1 and 2 years of eMDD use; 3 GPs were interviewed again after 10 months of use (these were from GP clinic 6 in Table 1). Both women and men participated in the individual and focus group interviews. The interviews lasted from half an hour to three-quarters of an hour, depending on the informants' information and schedules. The interview locations were conducted either in the GPs' offices or digitally via Skype for Business. Skype for Business was a solution to complete the data gathering after the lockdown restrictions that started in March 2020 due to COVID-19. Before each interview, the authors informed the participants about the project, and the GPs provided both written and verbal consent to participate in the study. The interviews were based on a semistructured interview guide, developed to obtain knowledge about issues related to the implementation and use of new technologies (eMDD). All the interviews began with an open question concerning the informant's experiences with eMDD. To ensure that the research question was covered, the interview guide was used as a checklist during the interviews [30]. In addition to the included questions in the guide, topics raised during the interviews were followed up, when appropriate, to obtain in-depth knowledge related to important issues for the GPs. The main topics in the interview guides concerned:

- How the doctors experienced using eMDD
- How the doctors experienced the start-up phase
- What changes occurred in the GPs' organization of clinical work
- What the doctors experienced as positive and as negative
- What the doctors felt could be improved

Table 1. Information about participants (N=26) and interview details.

GP[a] clinic number	Participants	GPs interviewed, n (%)	Interviews	Setting	Researcher
1	County 1	8 (30.8)	Focus group f2f[b]	Urban	TSB/EJ
2	County 2	1 (3.8)	Telephone interview	Rural	EJ
3	County 2	1 (3.8)	Telephone interview	Rural	EJ
4	County 2	1 (3.8)	Telephone interview	Rural	EJ
5	County 2	1 (3.8)	E-mail interview	Rural	EJ
6	County 3	8 (30.8)	Focus group f2f	Urban	TSB/EJ
7	County 3	1 (3.8)	Telephone interview	Urban	MKG
8	County 4	3 (11.5)	Individual interview f2f	Urban	MKG
9	County 4	1 (3.8)	Individual interview f2f	Urban	MKG
10	County 4	1 (3.8)	Individual interview f2f	Rural	MKG

[a]GP: general practitioner.

[b]f2f: face-to-face communication/meeting.

Analysis

The in-depth interviews were digitally recorded and then transcribed verbatim by a professional company. All the transcribed interviews were saved as files in NVivo (QSR International) to systemize the analysis [32]. Both authors were responsible for the interviews and the analysis of the material and also thoroughly discussed this several times during the analysis process. Both authors read all the interviews. To analyze the data, 4 steps of systematic text condensation were followed [30,33]. The authors first read all the interviews, initially to obtain a general impression and then to identify key themes. The authors read the interviews with special attention to the GPs' experiences during the start-up phase. NVivo was used to systematize relevant text [34] and then discussed and agreed on the key themes, categorized the text, and adjusted it, as needed. The categories were developed through an abductive and iterative process based on the topics in the interview guide [26]. The text was then condensed, analyzed, and discussed further, and finally merged into the revealed themes. The key themes are presented in the Results section, augmented by illustrative quotes.

We used a professional agency to translate from Norwegian to English.

Ethical Assessment

The project was approved by the data protection office (DPO) at the University Hospital of North-Norway (UNN; ethical approval no. 02003). The GPs involved in the study received both written and oral information about the study and were guaranteed anonymity before they agreed to voluntary participation. Information was given explaining that they could withdraw from the study at any time. The data are anonymized in the presentations.

Results

GPs' Experiences and Description of Using eMDD

In this section, the GPs' experiences and description of using eMDD are illustrated by presenting the findings, addressing the following research question: *How do GPs experience the introduction and use of e-prescribing for MDD?* The findings are represented by 4 emergent analytic themes: (1) start-up with eMDD ("Be patient"; (2) the need for training; (3) interaction, safety, and efficiency; and (4) the working day with eMDD.

Start-up With eMDD: "Be Patient"

Prior to the eMDD start-up, a joint introductory meeting was planned and held with each GP's office, with a video conference with pharmacies, the Norwegian Directorate of Health's IT department, and the GPs and technology suppliers. The objective of this meeting was to review how the introduction of eMDD was to be carried out technically and how the lists were to be submitted. The GPs described this meeting as useful, as was the manual sent by the pharmacy with a description of how the MDD patients should initially be registered. Registration was experienced as the most demanding process, and GPs responsible for several MDD patients reported a considerable uptick in work during the registration process. Shortly after the doctors signed up for the system as the responsible MDD GPs, they could start registering patients and medication lists. Several of the GPs emphasized the importance of having patience during the start-up process.

> *[You must] be patient with it. And be prepared that you may get a lot of messages from the pharmacy in the beginning. In the very stressful everyday life of general medicine, it can seem a bit like an extra stress in the beginning. But eventually, it becomes just part of the job, and then it becomes much easier. You also get a lot back and forth with the home service regarding medication. Be patient at first.*

Some GPs also experienced a number of technical problems when they had to register patients. The technology was slow at

start-up, and there was a considerable delay between when the information was added and when it was processed into the database. The GPs found it difficult to get started and described it as a slow system. This was revealed as a major obstacle in their everyday work, as it was not possible to do other tasks on the computers while the system was working on storing information. As a result, a number of GPs opted to spend time on this registration process in the evenings and on weekends.

The GPs' overall experience was that the registration took approximately 20-30 minutes per patient or MDD, reconciling the lists that were in different places (ie, at the pharmacy and with the doctor) and registering responsibility for the medication. Here, ensuring the coordination of lists in advance of the start-up proved to be an advantage. Lists from the pharmacy had been sent together with information about how to start with eMDD a few weeks before, which made it easier to complete the registration.

> *We first got a list from the pharmacies to know who our MDD patients were, and then we started the process of cleaning up medication lists and preparing. It was really like a pre-release, that. So, it was a long time in advance, so we had the opportunity to start working on it. But it still took a long time. However, you must go through each registration to see if it is correct. So, this took time anyway, even though all the work had been done in advance.*

According to some GPs, although there were only a few patients per doctor, they still experienced the job of converting to MDD e-prescriptions as an inconvenient but necessary task.

> *It is a bit of a job but not something that is unmanageable. What we have to think about is how the information should be provided, because since we have a pilot, we had meetings for this where we actually get paid to show up. That will not happen when this is rolled out, broadly.*

It was pointed out that the GPs who have the most patients on MDD will have a tremendous workload in the introductory phase but will probably also be the ones who get the most out of it once the system is up and running. One issue related to the introduction of eMDD was the uncertainty regarding the number of resources required to get started. As the GPs already felt overloaded with a considerable number of tasks in their everyday work life, it was a source of concern that there could be additional workload with a new system. To help achieve a smooth introduction, the GPs suggested that the workload associated with the introduction be made visible.

> *I think it's very good that when you start a project, it's just the beginning of a change that is for the better, and it's good to know what it entails. I would appreciate the project manager being honest and open and saying something about what to expect when we join.*

A clear expectation that emerged from the interviews was that eMDD would contribute to a simpler everyday work life for GPs and increase the quality and safety concerning the handling of medication. Several GPs emphasized that the benefit of such a system is safety: multidose drug prescriptions on paper and the use of fax for communication between different parties was described as a low-functional and old-fashioned system. They noted that the use of multidose drugs was often cumbersome and time-consuming and that it could be difficult to keep track of medications. As such, expectations for improvement were high.

> *So, a multidose is sometimes terrible, like that in paper form with lots of sources of error and lots of nonsense, faxes, and forms and triple lists and all that, so I ended up having to say no to MDD for new patients, because that scheme was low-grade quality. Getting an electronic multidose [system] has been welcome and something we have been waiting for.*

The motivation for several of the GPs was that any time spent getting started with the new system was something they would get back later.

The Need for Training

The GPs had different experiences concerning training to use the new system. Some GPs found that a letter from the pharmacy was sufficient for the start-up process, while others would have preferred to have attended courses. Those who were satisfied with the training using eMDD reported that having specific people (ie, at the pharmacy or technology supplier) to contact when they needed help with something was important. It was also easier for the GPs to become acquainted with getting started in doctor's offices where several others were interested in and had familiarized themselves with eMDD.

It was perceived as a problem by the GPs if there was no access to people who could be available as a resource (ie, those with a deeper understanding of the system). If a resource person was not available, this had a negative effect on implementation. Here, tasking someone with the role of a "superuser," who would get to know everything thoroughly and could be available as a resource for others, was recommended. Moreover, several of the GPs expressed that they had received neither the necessary information nor training and wished that someone had come to the office to introduce the project. A day-long conference with information and training with the project leaders was also recommended.

> *I wish they were clearer on the information about when it was to begin, and preferably, well, there are quite a few of us who are not too good at data (technology use), so to at least consider whether one can collect or create . . .*

The GPs pointed out that those who practiced in offices with fewer employees may have greater problems getting acquainted with new systems if they do not have a large patient base to register in the system. The GPs also pointed to training-related challenges as being rooted in a mixture of pedagogical shortcomings and some technical issues that made the initial workload heavier than it should have been. Here, the need for simple, quality training was highlighted—particularly training that could be undertaken during the workday rather than during the doctors' free time. In the interviews, it emerged that the GPs wanted to learn more about the system: for example, what the

pharmacy sees on the screen when the doctor sends something and what the hospital doctor can and should do with the medications being taken (ie, the patient's medication list). The GPs emphasized the necessity of adapting training related to their different needs, especially since there was variation in both their interest in and their desire for the digitalization of MDD. With regard to training, video clips and help from colleagues were highlighted as useful.

Interactions, Safety, and Efficiency

One of the most important tasks when first implementing eMDD is to clean up and update the medication lists with the correct information, and all the GPs described this as a tremendous undertaking. They also explained that it was important to approve the lists, be clear on how dosing takes place, and ensure that this is stated clearly; however, they experienced this as challenging when the system did not work properly.

> *It was a bit chaotic. We thought we had to delete old papers, and for some patients, there were huge lists of old prescriptions saved. But then we found out that it was possible to update without deleting old papers, and that made everything a little easier.*

The GPs described this process as quite labor intensive and, for many, surprising. Several related that they had not been mentally prepared for so much work, even though they had been informed well in advance to update the medication lists. One GP explained:

> *We had been sent what the pharmacy and the home service had on their lists, so it was up to date, and we thought it was mostly a push of a button. I'm not that computer savvy, and it took a lot of time. It was the use of time that was the problem [. . .] So, there was a lot of work then to start the process of getting an overview of all the lists.*

Another one shared:

> *Yes, so the advice is that you must always have an overview of your patients' medications and that you must enter and clean up the medication lists continuously. It must be "up to date."*

GPs described having to spend time cleaning up after hospital doctors who had prescribed new medications without deleting the valid prescriptions that were already in place, as the official regulations state that this is the GPs' responsibility. They, therefore, recommended that there must be an implementation period in which time and resources are set aside for training so that everyone understands the importance of doing this.

GPs described both positive and negative experiences related to interaction and safety when using eMDD. For example, the nursing and care e-messages between GPs and the home care staff were experienced as smart and were defined as a "safety valve" concerning communicating changes in the medication lists. The GPs felt this provided a better overview for all actors with regard to determining the correct medication. However, communicating changes in the medication list to the multidose pharmacy was experienced as more uncertain:

> *I'm not always quite sure if they got it. It has been—or we have to write physically as a message at the bottom, "I have changed so and so." I have actually experienced that they have not done exactly as I have said.*

This challenge was explained as being partly due to a lack of knowledge regarding what the technical aspect looks like at the pharmacy (ie, whether it is physically possible for the GP to make a mistake when sending an MDD list to them). The question is whether it goes to a machine and the machine makes all the mistakes or whether it is the case that a person is responsible for what is to be in each small bag.

Several points emerged in the interviews related to weaknesses in the safety of medication use for patients. A problem highlighted by all involved parts in the medication chain was that there was a big safety gap related to the fact that medications prescribed with e-prescriptions can be picked up twice. The GPs pointed to an example: if an electronic prescription is legal for a year, the system is not structured so that it is locked in the multidose drug list. This means that the patient can pick up the medication by themselves, regardless of what is packaged in the MDD. The pharmacy should be able to determine that the prescriptions have all been picked up, but instead, they are packaged in the MDD, and the user will get double medication. To increase safety and overview for all parties, the GPs thus stated that it is important to emphasize a thorough review of medication lists, structured as part of their everyday workday.

> *The important thing was to have updated medication lists, that we had to make sure that we did not have any magistral prescriptions, or any reminders we had to ourselves, or that there were messages to the home care staff in the medication lists. Because there were some things we did before to make things work, which do not work at all if you have e-multidose. So, it took a while to clean up those lists.*

The Working Day With eMDD

The GPs also had different experiences around the use of eMDD in their everyday workday—this seemed related to whether there were clear lines of communication between all involved parties. One frustration noted by many of the GPs was that changes made to the system were not always registered. They would then receive a message from the pharmacy to discontinue and recall the medication list and prescribe again. One GP shared this experience:

> *Sometimes, I have tried to discontinue medication 5 times and yet it has not worked. Then the message comes back from the pharmacy, and they write smiley faces and try to be nice to us, and say we are sorry, but you actually have to stop again and prescribe again.*

One GP suggested a phone-a-friend approach as a solution to this issue.

> *There are programs on TV that have an option called "phone a friend." And you can at least do that at least once, so I can tell you that you are allowed to call me*

> *in the evening, but sitting together and watching it together, I think that might have solved his frustration and your problem, so probably everyone would have saved time.*

Many GPs reported that after the initial start-up process, once the system had been in use for a while, it facilitated a better working day. As one stated,

> *Errors in the lists—they are not there anymore. So now there is a good flow in our workflow, so it is an integral part of our everyday life that we do not think so much about anymore.*

Another GP shared his experience:

> *I am very happy with e-multidose, it is very good. We can reduce the use of paper—as long as it works, it is absolutely fantastic. So, it's just to make it work, but lately it has been very smooth, so there have been no problems in recent weeks, and very few messages from the home nurse and from the pharmacy. When things are established, it rolls smoothly.*

Even if it was a challenge during the implementation phase, most of the GPs welcomed the eMDD:

> *I think no one really knew what they were getting into, so everyone was optimistic and looking forward to finally dropping the fax and stuff.*

Discussion

Principal Findings

In the Results section, we presented GPs' experiences with implementing eMDD, focusing on GPs' information and training needs and their experiences with the start-up process, including the coordination of lists, safety and effectiveness, and changes to their working day [23]. There are variations in different health providers' and GPs' needs regarding training and information and considerable variation in competence and motivation related to the use of digital tools [35]. There were also different degrees of understanding concerning the everyday work of the other actors in the medication chain. In particular, the harmonization of medication lists related to the use of time, expenditures, and challenges with technological solutions in the introduction phase was emphasized as a challenge [36,37]. Overall, GPs who have started using the system report great benefits; these are largely related to an increased overview of patients' total medication lists, less time spent on prescribing prescriptions, and increased collaboration with pharmacies and nurses, both in service from providers in homes and in nursing homes.

Previous studies have shown that better availability of patients' overall medication increases patient safety and increases collaboration between different health care providers. In addition, access to a patient's medication list and health information enhances safety and saves GPs time [38]. One reason is the faster updating of prescriptions electronically. One of the most positive things about eMDD from the GPs' perspective, compared to the use of paper and fax, is a better overview of lists and that prescribing can be done immediately and increases the chances of the information arriving [22]. As such, to achieve quality implementation, it is important to develop systems that ensure quality information and training provision at start-up; it is equally important to have quality guidelines in place and technology that promotes interaction between all involved parties and ensures safety for patients [39]. To obtain this, it is important to gain a complete and accurate overview of each patient's medication needs. It is essential that the type of medication and dosage be included in the medication list—this, in turn, ensures professional justification and enhances both the quality of services and the patient's quality of life.

As revealed in the Results section, however, there are still several challenges associated with this. Regarding organization and collaboration, the GPs reported a lack of knowledge about what the medication chain looks like for each individual involved—a source of concern as they felt this could affect both safety and effectiveness with regard to medication management. Another challenge was the delay experienced between when information was added to the patients' medication list and when the system reflected the updates. There are several possible solutions to this issue. This could be an opportunity for hospital doctors to discontinue a medication that should be removed from the MDD list. Increased communication and understanding of deadlines between the various actors, and a good support service related to the digital system(s), such as the technology provider, may ensure right medication. Digitalization helps ensure faster and more secure transfer of information when the technology is implemented in an appropriate way [40]. Research has shown that both GPs and employees in the home care service experience MDD as contributing to quality improvement related to patient overview and safety when patients are taking multiple medications [21]. In this study, the primary attitudes toward eMDD among the GPs were positive; they felt it facilitated better patient safety and was efficient and professionally justified. However, they also emphasized that to create and implement a well-functioning eMDD solution, collaboration between all actors is required. The question remains whether the use of eMDD actually contributes to the realization of gains via increased efficiency (ie, through reducing time spent on prescribing and improving interaction and patient safety). The process of getting started with eMDD was labor intensive for the GPs. However, once they spent the time necessary to establish an updated and correct medication list for each MDD patient, it proved a time saver during their everyday workday and contributed to increased patient safety [36]. Nevertheless, there will always be variations within and between municipalities, GPs, and the specialist service; as such, using eMDD on a broader scale in Norway and other countries will necessitate a focus on ensuring that the digital solutions are implemented with quality information, training and structure, and standardized solutions in place as far as possible.

Implications for Practice

There are some important issues to follow up on in the introductory phase and the scaling-up process of introducing eMDD in GPs' offices. Several of the GPs in this study looked forward to the project's start-up, yet many pointed out that it might be difficult to handle the extra tasks. To sum up, good routines are necessary for training all stakeholders, including GPs, pharmacies, municipal health services, technology

suppliers, and patients, where appropriate. Making the appointment of a superuser responsible for eMDD, who can follow up when needed, is also of high importance. It was especially pointed out that a specific contact person for both GPs and nurses at pharmacies when complications occur will ease the implementation [21]. The GPs felt that increased contact and collaboration with the pharmacy could have helped simplify the work. Clear placement of responsibility for solving challenges that arise is also needed; this also applies to support for technology challenges. There must be a provision of extra time to register everything correctly when starting eMDD. Having to do the same task repeatedly was time-consuming and was noted as potentially hindering the GPs' ability—and willingness—to implement eMDD as part of their everyday work lives. eMDD seems like a safer and more effective solution when implemented in the organization for all the included parts. Nevertheless, these data may contribute to a greater reflection on—and discussion about—the current, rapid implementation of electronic prescription of medication in the health services and the challenges that may appear.

Limitations/Weaknesses of the Study and Issues for Further Research

The study was performed during the implementation of eMDD and followed up with a few interviews after 3-6 months to explore experiences with the start-up process of eMDD by the GPs. A potential weakness of the study is its reliance on both physical and digital interviews with GPs. As such, the information derived from the interviews may have been different if the interviews had been conducted in person. Another weakness is related to the use of different interview strategies; however, this may also have strengthened the analysis by investigating both individual opinions and opinions reflected in a group of GPs. We further acknowledge the variation of the GPs' experience with eMDD, with some of them being experienced only for 2 months and others for 2 years, as a limitation.

Conclusion

The literature on the topic is growing but still limited, and more research is needed to develop digital prescription of medication to enhance safety for all included parts, especially the users. Awareness of the hindrances revealed both in earlier research and in this study may strengthen the motivation and establish routines including stakeholders and support from both the pharmacies and the technology provider for launching the digital solution eMDD as a working tool for GPs. There is a need for further investigation, including qualitative research, to build solid and evidence-based knowledge that can contribute to developing tailored handling of medication for multidose drug users. Further research should focus on service users' experiences, cocreation between different stakeholders, and how to scale up the use of eMDD, while ensuring that the use of eMDD is appropriate, safe, and available for end users (patient), next of kin, and health service providers (eg, GPs, pharmacists, and nurses).

Acknowledgments

The authors would like to thank all general practitioners (GPs) who participated in this study and Elin Johnsen for her contribution to the data collection. This study was internally funded by the Norwegian Centre for E-health Research.

Authors' Contributions

Both authors made significant contributions to the manuscript. The study was conceived by MKG and TSB and was drafted in close cooperation. Both MKG and TSB collected data and contributed to the analysis. The researcher EJ also collected data. The manuscript was written by MKG and TSB, and both authors read and approved of the final manuscript.

Conflicts of Interest

None declared.

References

1. Norwegian Directorate of Health. Multidose in E-Prescription. Oslo: Norwegian Directorate of Health; Feb 12, 2021.
2. Ministry of Health and Care Service. National Health and Hospital Plan 2020-2023. Oslo: Ministry of Health and Care Services; Nov 22, 2019.
3. Oktora MP, Denig P, Bos JHJ, Schuiling-Veninga CCM, Hak E. Trends in polypharmacy and dispensed drugs among adults in the Netherlands as compared to the United States. PLoS One 2019 Mar 22;14(3):e0214240 e021424 [FREE Full text] [doi: 10.1371/journal.pone.0214240] [Medline: 30901377]
4. Fastbom J, Johnell K. National indicators for quality of drug therapy in older persons: the Swedish experience from the first 10 years. Drugs Aging 2015 Mar 12;32(3):189-199 [FREE Full text] [doi: 10.1007/s40266-015-0242-4] [Medline: 25673123]
5. Hermann M, Carstens N, Kvinge L, Fjell A, Wennersberg M, Folleso K, et al. Polypharmacy and potential drug–drug interactions in home-dwelling older people: a cross-sectional study. J Multidiscip Healthc 2021 Mar;Volume 14:589-597. [doi: 10.2147/jmdh.s297423]

6. Fog AF, Straand J, Engedal K, Blix HS. Drug use differs by care level. A cross-sectional comparison between older people living at home or in a nursing home in Oslo, Norway. BMC Geriatr 2019 Feb 19;19(1):49 [FREE Full text] [doi: 10.1186/s12877-019-1064-8] [Medline: 30782115]
7. Wekre LJ, Bakken K, Garåsen H, Grimsmo A. GPs' prescription routines and cooperation with other healthcare personnel before and after implementation of multidose drug dispensing. Scand J Public Health 2012 Aug 07;40(6):523-530. [doi: 10.1177/1403494812455468] [Medline: 22872598]
8. Frøyland H. Legers synspunkter på multidosepakkede legemidler. Norsk Farmaceutisk Tidsskrift 2012;4:19-21.
9. Josendal AV, Bergmo TS, Granas AG. Potentially inappropriate prescribing to older patients receiving multidose drug dispensing. BMC Geriatr 2020 Aug 05;20(1):272 [FREE Full text] [doi: 10.1186/s12877-020-01665-x] [Medline: 32758129]
10. Jøsendal A. Multidose i e-resept En sammenligning av avvik i legemiddellistene (LIB) mellom fastleger og apotek før og etter innføring av multidose i e-resept (SERIE NSE-report ISSN 2535-2571). Tromsø: Norwegian Center for E-Health Research; 2017:01-2017.
11. Sinnemäki J, Airaksinen M, Valaste M, Saastamoinen LK. Impact of the automated dose dispensing with medication review on geriatric primary care patients drug use in Finland: a nationwide cohort study with matched controls. Scand J Prim Health Care 2017 Dec 10;35(4):379-386 [FREE Full text] [doi: 10.1080/02813432.2017.1398933] [Medline: 29125004]
12. Bardage C, Ekedahl A, Ring L. Health care professionals' perspectives on automated multi-dose drug dispensing. Pharm Pract (Granada) 2014 Oct;12(4):470 [FREE Full text] [doi: 10.4321/s1886-36552014000400005] [Medline: 25580170]
13. Holbø K, Das A, Bøthun S, Formanek M, Halvorsen T. Multidose service for home dwellers: the users' experiences and a need for new solutions. Nord Welfare Res 2019 Mar 20;4(01):20-31 [FREE Full text]
14. Stewart D, Mair A, Wilson M, Kardas P, Lewek P, Alonso A, SIMPATHY Consortium. Guidance to manage inappropriate polypharmacy in older people: systematic review and future developments. Expert Opin Drug Saf 2017 Feb 25;16(2):203-213. [doi: 10.1080/14740338.2017.1265503] [Medline: 27885844]
15. Lester CA, Tu L, Ding Y, Flynn AJ. Detecting potential medication selection errors during outpatient pharmacy processing of electronic prescriptions with the RxNorm Application Programming Interface: retrospective observational cohort study. JMIR Med Inform 2020 Mar 11;8(3):e16073 [FREE Full text] [doi: 10.2196/16073] [Medline: 32044760]
16. Wekre LJ, Spigset O, Sletvold O, Sund JK, Grimsmo A. Multidose drug dispensing and discrepancies between medication records. Qual Saf Health Care 2010 Oct 04;19(5):e42-e42. [doi: 10.1136/qshc.2009.038745] [Medline: 20688757]
17. Johnsen E, Jøsendal AV, Bergmo TS. The e-multidose is better for patients' safety than dosette boxes and faxes. Sykepleien 2018:e69983. [doi: 10.4220/sykepleienf.2018.69983en]
18. Norwegian Pharmacy Association. Pharmacies and Pharmaceuticals in Norway. Oslo: Norwegian Pharmacy Association; Jan 01, 2021.
19. Alshahrani F, Marriott JF, Cox AR. A qualitative study of prescribing errors among multi-professional prescribers within an e-prescribing system. Int J Clin Pharm 2021 Aug 09;43(4):884-892 [FREE Full text] [doi: 10.1007/s11096-020-01192-0] [Medline: 33165835]
20. Hjelle K. Multidose: Status og veien videre. 2015 Jun 01. URL: https://tinyurl.com/24un4xe9 [accessed 2021-01-12]
21. Bergmo T, Jøsendal A, Johnsen E. Factors easing the transition from paper to electronic prescribing. Linkoping: Linköping Electronic Conference Proceedings ISSN 1650-3686 e-ISSN 1650-3740; 2019 Nov 12 Presented at: SHI 2019 Proceedings of the 17th Scandinavian Conference on Health Informatics; November 12-13, 2019; Oslo, Norway.
22. Jøsendal A, Bergmo T. How discrepancies in medication records affect the creation and trust in a shared electronic medication list in Norway. 2019 Nov 12 Presented at: SHI 2019 Proceedings of the 17th Scandinavian Conference on Health Informatics; November 12-13, 2019; Oslo, Norway.
23. Garfield S, Furniss D, Husson F, Etkind M, Williams M, Norton J, et al. How can patient-held lists of medication enhance patient safety? A mixed-methods study with a focus on user experience. BMJ Qual Saf 2020 Sep 16;29(9):764-773 [FREE Full text] [doi: 10.1136/bmjqs-2019-010194] [Medline: 31949006]
24. Grossman J, Cross D, Boukus E, Cohen G. Transmitting and processing electronic prescriptions: experiences of physician practices and pharmacies. J Am Med Inform Assoc 2012;19(3):353-359 [FREE Full text] [doi: 10.1136/amiajnl-2011-000515] [Medline: 22101907]
25. Huber T, Brinkmann F, Lim S, Schröder C, Stekhoven DJ, Marti WR, et al. Implementation of an IT-guided checklist to improve the quality of medication history records at hospital admission. Int J Clin Pharm 2017 Dec 29;39(6):1312-1319 [FREE Full text] [doi: 10.1007/s11096-017-0545-0] [Medline: 29082460]
26. Blaikie N. Approaches to Social Enquiry: Advancing Knowledge. Cambridge: Polity Press; 2007.
27. Peirce CS. Pragmatism as a Principle and Method of Right Thinking: The 1903 Harvard Lectures on Pragmatism. Albany, NY: SUNY Press; 1997.
28. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 08;320(7227):114-116 [FREE Full text] [doi: 10.1136/bmj.320.7227.114] [Medline: 10625273]
29. Kvale S, Brinkmann S. Interviews: Learning the Craft of Qualitative Research Interviewing. "Det kvalitative forskningsintervju". London: SAGE; 2009.
30. Kvale S, Brinkmann S. Doing interviews. In: Conducting an Craft of Qualitative Research Interviewing. "Det kvalitative forskningsintervju". London: SAGE; 2015.

31. Alvesson M, Sköldberg K. Reflexive Methodology: New Vistas for Qualitative Research. London: SAGE; 2017.
32. AlYahmady HH, Al Abri SS. Using Nvivo for data analysis in qualitative research. IIJE 2013 Feb;2(2):181-186. [doi: 10.12816/0002914]
33. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec 04;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
34. Maher C, Hadfield M, Hutchings M, de Eyto A. Ensuring rigor in qualitative data analysis. Int J Qual Methods 2018 Jul 10;17(1):160940691878636. [doi: 10.1177/1609406918786362]
35. Sarradon-Eck A, Bouchez T, Auroy L, Schuers M, Darmon D. Attitudes of general practitioners toward prescription of mobile health apps: qualitative study. JMIR Mhealth Uhealth 2021 Mar 04;9(3):e21795 [FREE Full text] [doi: 10.2196/21795] [Medline: 33661123]
36. Manskow US, Kristiansen TT. Challenges faced by health professionals in obtaining correct medication information in the absence of a shared digital medication list. Pharmacy (Basel) 2021 Feb 22;9(1):46 [FREE Full text] [doi: 10.3390/pharmacy9010046] [Medline: 33671820]
37. Kleiven HH, Ljunggren B, Solbjør M. Health professionals' experiences with the implementation of a digital medication dispenser in home care services: a qualitative study. BMC Health Serv Res 2020 Apr 16;20(1):320 [FREE Full text] [doi: 10.1186/s12913-020-05191-9] [Medline: 32299431]
38. Josendal AV, Bergmo TS. From paper to e-prescribing of multidose drug dispensing: a qualitative study of workflow in a community care setting. Pharmacy (Basel) 2021 Feb 16;9(1):41 [FREE Full text] [doi: 10.3390/pharmacy9010041] [Medline: 33669475]
39. Nilsen ER, Dugstad J, Eide H, Gullslett MK, Eide T. Exploring resistance to implementation of welfare technology in municipal healthcare services: a longitudinal case study. BMC Health Serv Res 2016 Nov 15;16(1):657 [FREE Full text] [doi: 10.1186/s12913-016-1913-5] [Medline: 27846834]
40. 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 Apr 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]

Abbreviations

eMDD: E-prescribing of multidose drug dispensing
GP: general practitioner

Edited by A Kushniruk; submitted 25.01.21; peer-reviewed by H Kauppinen, H Imeri, H De Loof, E McCourt, G Carot-Sans; comments to author 28.02.21; revised version received 21.05.21; accepted 27.11.21; published 17.01.22.

Please cite as:
Gullslett MK, Strand Bergmo T
Implementation of E-prescription for Multidose Dispensed Drugs: Qualitative Study of General Practitioners' Experiences
JMIR Hum Factors 2022;9(1):e27431
URL: https://humanfactors.jmir.org/2022/1/e27431
doi:10.2196/27431
PMID:35037881

Original Paper

Identifying Ethical and Culturally Responsive Research Activities to Build Trust and Improve Participation of Black Sexual Minority Men in Pre-Exposure Prophylaxis Telehealth Clinical Trials: Qualitative Study

Derek T Dangerfield II[1,2], PhD; Charleen Wylie[1]

[1]Johns Hopkins School of Nursing, Baltimore, MD, United States

[2]Us Helping Us, People Into Living, Inc, Washington, DC, United States

Corresponding Author:
Derek T Dangerfield II, PhD
Johns Hopkins School of Nursing
525 N Wolfe St
Baltimore, MD, 21205
United States
Phone: 1 4439839283
Email: ddanger2@jhu.edu

Abstract

Background: Telehealth interventions could improve pre-exposure prophylaxis (PrEP) initiation and adherence in high HIV incidence groups such as young Black sexual minority men (BSMM). However, young BSMM remain distrustful of and underrepresented in clinical trials. Therefore, ethical and culturally responsive ways are needed to build trust and improve their participation in PrEP telehealth clinical trials.

Objective: To bridge this gap, this study identified ethical and culturally responsive activities to build trust and improve participation of young BSMM in PrEP telehealth clinical trials.

Methods: We obtained data from 7 virtual, synchronous focus groups that were conducted from April to August 2020 and consisted of 28 BSMM aged 18-34 years. Focus groups included a brief survey distributed online via Qualtrics followed by a virtual, synchronous focus group conducted via Zoom that lasted between 50 and 75 minutes. Focus groups were stratified by age (18- to 24-year-old participants and 25- to 34-year-old participants), outlined the components of an example PrEP telehealth randomized controlled trial, and included questions on domains of the study design—research motivations, study funding, recruitment activities, informed consent details, randomization, follow-up, and end of study activities. Participants were asked targeted questions regarding the ethics and trustworthiness of the study and ways in which researchers could gain their trust through the protocol used in the PrEP telehealth clinical trial.

Results: The focus groups included 2 groups of 18- to 24-year-old participants and 5 groups of 25- to 34-year-old participants. The mean age of participants was 27.2 years (SD 4.4 years). Of the 28 participants, 10 (36%) reported a bachelor's degree to be their highest completed education level and 6 (21%) reported some graduate degree or higher to be their highest completed education level. Most participants (16/28, 57%) reported that they worked full-time and that they were single or not in a committed relationship (21/28, 75%). Most participants (24/28, 86%) reported that they used at least one drug before sex in the 6 months prior to the study. All participants reported that they heard about PrEP and 36% (10/28) were current PrEP users. Overall, the focus groups yielded themes related to the impact of researcher intentions, study funding, recruitment activities, informed consent details, randomization, and study team interactions during and after the study on trust and participation in the clinical trial.

Conclusions: Medical and research mistrust persists among BSMM. This study identified several ethical and culturally responsive activities to build trust and improve participation of young BSMM in PrEP telehealth clinical trials. Future studies should assess the relative impact of implementing these findings on research participation in a PrEP telehealth clinical trial.

(JMIR Hum Factors 2022;9(1):e28798) doi:10.2196/28798

KEYWORDS

HIV; sexual health; stigma; medical mistrust; PrEP; telehealth; medication adherence; minorities; focus groups; sexual minorities; mobile phone

Introduction

To reduce disparity in the United States, it is crucial to improve participation of young Black gay, Black bisexual, and other Black sexual minority men (BSMM) in HIV pre-exposure prophylaxis (PrEP) telehealth clinical trials [1,2]. Without substantial improvement in prevention activities, BSMM have an estimated 50% lifetime risk of HIV [3]. Between 2010 and 2017, HIV incidence increased by 42% among BSMM aged 25-34 years [4]. In 2018, HIV infection in BSMM accounted for 26% of all HIV infections in gay and bisexual men in the United States; approximately 75% of newly diagnosed HIV infections in BSMM were in those under the age of 35 years [4]. Data show that PrEP substantially reduces the risk of HIV infection [5-7] and that telehealth interventions could improve PrEP initiation and adherence [8]. However, young BSMM remain underrepresented in PrEP clinical trials [9]. Increased participation in clinical trials is needed from BSMM to improve PrEP telehealth protocols.

Telehealth refers to the use of telecommunication technology to support long-distance clinical health care, health education, and health administration [10,11]. Telehealth programs are conducted remotely by a clinician via applications that are accessible on a smartphone, tablet, or computer through a video, a telephone call, or an SMS text messaging platform compliant with the Health Insurance Portability and Accountability Act of 1996 [8,11,12]. Currently, standard PrEP clinical care requires in-person visits with a clinician along with laboratory testing for HIV, sexually transmitted infections, hepatitis B, and serum creatinine levels prior to prescription [13]. However, telehealth protocols allow PrEP patients to communicate virtually with clinicians and then visit a local outpatient clinic, laboratory, or public health facility for testing [12]. Some telehealth studies mail at-home self-testing kits to patients and patients then need to return the specimens for testing prior to PrEP prescription [14]. Telehealth protocols could overcome some of the structural barriers to standard PrEP care, such as limited transportation, anticipated stigma in health care settings, and privacy concerns [15-17]. However, it has been challenging to engage young BSMM in PrEP telehealth clinical trials in part because of medical and research mistrust and the traditionally experienced stigma, discrimination in health care settings, and competing socioeconomic demands [15,18].

Medical mistrust refers to the lack of trust in the motives of and treatment by individuals and organizations associated with health care institutions [19,20]. Medical mistrust is attributed to historically limited health care access for Black Americans, including young BSMM, and maltreatment of this population by health care professionals and medical researchers [18,21]. Examples of maltreatment of Black Americans by health care professionals and medical researchers include discrimination, treatment refusal, treatment deception, and enrollment of unwilling participants in clinical experiments for research [15,21,22]. Many of these medical and research activities were supported by US racial segregation laws and a lack of established ethical research and medical guidelines [22]. The history of research and medical institutional mistreatment of racial or ethnic minority populations has impacted care satisfaction, treatment adherence, and clinical research participation among Black Americans, including young BSMM [19,23,24]. Medical mistrust is salient for BSMM [18,25,26]. Many studies have documented challenges in engaging young BSMM in HIV-prevention clinical trial research generally because of medical mistrust [24,27,28]. However, more information is needed to identify ways to improve trust and participation of this vulnerable population in PrEP telehealth clinical trial research.

To bridge this gap, this study identified ethical and culturally responsive study activities to improve participation of BSMM in PrEP telehealth clinical trials. History-based models of trust [29,30] provide a useful framework to guide this study. These models posit that trust results from cumulated, actual, or vicarious experiences and that cumulative negative experiences in society and health care settings disrupt one's sense of safety [30]. Therefore, mistrust can increase as a safety mechanism. Although studies have identified the role of medical mistrust among BSMM in PrEP research engagement [15,18,31], more targeted data are needed to build trust, reduce mistrust, and improve participation in PrEP clinical trials for this group. Little is known about specific ways to improve trust in this population along the research process, that is, during recruitment, while obtaining informed consent, and at enrollment, follow-up, and study completion. The findings of this study could be used to design an ethical and culturally responsive PrEP telehealth intervention for BSMM.

Methods

Study Recruitment and Participants

We collected data from 7 virtual, synchronous focus groups that were conducted from April to August 2020 and consisted of 28 BSMM aged 18-34 years. Individuals were recruited using a combination of active and passive strategies. Active recruitment included contacting participants from other research studies who provided written consent to be contacted for future research. Passive recruitment included advertising the study on Craigslist and obtaining referrals from participants in the study. Eligibility for participation in the study was defined by the following criteria: self-identification as a Black or African American man, age between 18 and 34 years, self-report of HIV-negative status, report of oral or anal intercourse with at least one male partner in the previous 6 months, or self-identification as a gay, bisexual, queer, or non-heterosexual individual.

Data Collection

Because of COVID-19, data collection for the focus groups was updated to a virtual, synchronous format for safety. Details of

the protocol for conducting the virtual, synchronous focus group have been published [32]. Focus groups included a brief survey distributed online via Qualtrics followed by a virtual, synchronous focus group conducted via Zoom that lasted between 50 and 75 minutes. Focus groups were led by 2 experienced facilitators and were stratified by age (18- to 24-year-old participants and 25- to 34-year-old participants). One facilitator conducted the groups and recorded notes and the other scheduled the groups, recorded field notes, observed group dynamics, provided technical support for participants who had difficulty connecting to the meeting (eg, use of the wrong password), and confirmed time and attendance [32]. For data on ethical and culturally responsive ways to build trust and improve participation of BSMM in PrEP telehealth clinical trials, the focus groups outlined the components of an example PrEP telehealth randomized controlled trial and included questions on domains of the study design—research motivations, study funding, recruitment activities, informed consent details, randomization, follow-up, and end of study activities. Participants were asked targeted questions regarding the ethics and trustworthiness of the study and ways in which researchers could gain their trust through the protocol used in the PrEP telemedicine trial. Participants were given a $75 electronic gift card as compensation for their participation. All participants provided oral informed consent that was documented by the study team prior to beginning the focus group [32]. All study procedures were approved by the Johns Hopkins School of Medicine Institutional Review Board.

Qualitative Data Analysis

Virtual, synchronous focus groups were audiorecorded using a handheld digital recorder to increase anonymity among the participants [32] and files were transcribed by a private, institutional review board–approved transcription service. The focus group facilitators reviewed all focus group transcripts and notes and developed a codebook for descriptive thematic analysis using Atlas.ti 8.4 (ATLAS.ti Scientific Software Development GmbH). Themes were identified using an adapted "pile sorting approach" [33,34]. Specifically, all quotes that were associated with specific codes in Atlas.ti 8.4 were electronically copied and pasted into an Excel sheet and organized by code. Quotes were reviewed by the lead investigator and sorted into "piles" for similarity within the Excel sheet. These piles represented the themes associated with specific focus group questions and codes. Themes were identified as patterns that were associated with specific focus group questions or expressions that provided novel responses to domains within the focus group guide [33,35,36]. To identify a range of themes, novel responses by at least one person in the group were also considered [37]. Between-group analysis was also conducted to identify potential differences in themes by age. Data presented in this study represent the range of themes related to culturally responsive ways to build trust and improve participation of BSMM in PrEP telehealth clinical trials.

Results

Demographic Characteristics

The focus groups included 2 groups of 18- to 24-year-old participants and 5 groups of 25- to 34-year-old participants. The mean age of the participants was 27.2 years (SD 4.4 years). Of the 28 participants, 10 (36%) reported a bachelor's degree to be their highest completed education level and 6 (21%) reported some graduate degree or higher to be their highest completed education level. Most participants (16/28, 57%) reported that they worked full-time and that they were single or not in a committed relationship (21/28, 75%). Most participants (24/28, 86%) reported that they used at least one drug before sex in the 6 months before the study. All participants reported that they heard about PrEP and 36% (10/28) were current PrEP users (Table 1).

Table 1. Sociodemographic and behavioral characteristics of focus group participants (N=28).

Characteristics	Value
Age range (years)	19-34
Age (years), mean (SD)	27.2 (4)
Sexual orientation, n (%)	
Homosexual, gay, or same gender loving	26 (92)
Bisexual	2 (7)
Highest education completed, n (%)	
Grade 11 or less	1 (3)
Grade 12 or GED[a] equivalent	5 (17)
Some college	4 (14)
Bachelor's degree	10 (36)
Some graduate degree or more	6 (21)
Employment status, n (%)	
Unemployed or not working	4 (3)
Part-time	3 (10)
Full-time	16 (57)
Other	3 (10)
Marital status, n (%)	
Single or not in a committed relationship	21 (75)
In a committed relationship	5 (17)
Married	2 (7)
Annual income, n (%)	
Less than $20,000	6 (21)
$20,000-$30,000	3 (10)
$30,000-$40,000	2 (7)
$40,000-$50,000	5 (17)
More than $50,000	10 (36)
Drugs used before sex in the past 6 months, n (%)	
Marijuana	24 (85)
Poppers	9 (32)
Ecstasy	5 (17)
Powder cocaine	5 (17)
Prescription painkillers	1 (3)
Ever tested for HIV, n (%)	26 (92)
Ever heard of PrEP[b], n (%)	28 (100)
Ever used PrEP, n (%)	15 (53)
Currently using PrEP, n (%)	10 (36)
Interested in PrEP injectable, n (%)	21 (75)
Interested in PrEP telemedicine, n (%)	
Yes	20 (71)
No	6 (21)
Don't know	2 (7)

Characteristics	Value
Interested in sexually transmitted infection (syphilis) PrEP, n (%)	16 (57)

[a]GED: General Education Development.

[b]PrEP: pre-exposure prophylaxis.

Thematic Findings

Overall, the focus groups yielded themes regarding the culturally responsive ways in which researchers could build trust and improve participation in PrEP telehealth clinical trial research. Specifically, participants shared comments on how researcher intentions, study funding, recruitment activities, informed consent details, randomization, and study team interactions during and after the study could impact trust and participation. Themes regarding each domain along the exemplar clinical trial research protocol are explained.

Intentions Behind Study Funding

Overall, participants had mixed feelings about trusting PrEP clinical trial research that was funded by pharmaceutical companies and the US government. They perceived both the government and pharmaceutical companies to be more invested in the profitability of PrEP dissemination than in the promotion of health for BSMM. They felt that an "investment in positive outcomes" could ensure greater participant safety, but they could also experience less safety because of funders' lack of "care" for positive outcomes among Black patients. One group of 25- to 34-year-old participants shared that they had greater trust in studies funded by private foundations because of the perception that they could more easily hold these foundations accountable for any adverse events in the protocol than the government or a pharmaceutical company. Participants across groups shared that trust in the safety of the telemedicine research protocol could be gained with knowledge of the intentions of the funding source.

> *What do y'all think about a study like this being funded by a pharmaceutical company, versus the government, versus a private foundation?* [Facilitator, 17:58]
>
> *It's all the same damn thing to me.* [P4, 18:09]
>
> *--I feel like pharmaceutical companies have more, you know, interest of duty kind of studies because these studies directly benefit them. And so that might make space for them to become more exploitive. Because, you know, they get something out of these studies directly versus the government. I don't know if the money train hits them the same way as a pharmaceutical company. So the government might be less inclined promote or, organize, a very exploitive study versus pharmaceutical companies, I think.* [P3, 18:31]

Another group of 25- to 34-year-old participants shared the following comments:

> *Like, there's money being made. And when money's being made like that it becomes ulterior agenda. Like, even if it was something that was positive to start, it's like, okay, well, there's money in the research. There's money in getting out there. There's money in finding another way to do it. It's--this is being pushed by money. Like, I mean, it could be a health thing. But it's a health thing being pushed by money and money. So I think that this is still a corporation... Even if you already got it [HIV], I got a pill for that too. So it's like, no, I don't trust the shit, right. I don't trust any of it anymore. And it's not that I don't trust the research behind it. I don't trust the way the research is being presented.* [P4, 20:59]
>
> *… I'm not stupid, like, I understand there's a business side to all of this shit. So I'm just wondering what the--even if the true intent is to help people, I'm wondering what the gag is. I'm waiting for the shoe to drop.* [P2, 22:31]
>
> *To that point both Green and Red's point. Y'all got to be making a shitload of money that you can just give out $75 to us for a survey. It's a nominal cost…And to Red's standpoint, can there be good intentions behind it? Yes. Could the pill actually be working? Yes. But it's kind of like, at what cost? What am I really giving away?* [P3, 23:31]
>
> *Yeah. Yeah. Give a piece of your liver just in case you get drunk and fuck.* [P4, 25:13]

These sentiments suggest a dissonance in attempts to identify altruistic intentions of funding PrEP clinical trials relative to the anticipation of maltreatment because of the profitability of positive outcomes.

Study Recruitment: Having Black Investigative Teams and "Care"

Participants across groups shared that they would have greater trust and interest in the PrEP telehealth clinical trial if the study was led by Black researchers. They mentioned that Black researchers would provide better care during the study and interpret findings better than non-Black researchers because the participants believed that Black researchers had greater investment in the overall wellness of BSMM. When asked how researchers could gain the trust of BSMM to increase their participation in PrEP clinical trial research, participants in a 25- to 34-year-old group agreed with the man who said,

> *Well, I think we need to see more black queer men doing these research studies, and not necessarily the face of it but--but actually running them from the start also. So don't just put a black face on there for, you know, image purposes or, you know, recruitment purposes, but actually have someone like us running the whole damn thing.* [P4, 01:09:05]
>
> *Why?* [Facilitator, 01:09:34]
>
> *Well, like I said, going back to like I was saying before, you know, in order for me to trust you, I have*

> *to feel like you have, uh, walked similar spaces that I have. So therefore, you can--you understand where I'm coming from when I'm telling you all this information about it.* [P4, 01:09:37]

When asked how non-Black researchers could gain trust, every group mentioned the need for non-Black researchers to be "involved in the community" and collaborate with community-based organizations, including the 2 groups that suggested that the race of the investigative team did not matter if they perceived that they "cared about us." However, 1 person in a 25- to 34-year-old group said,

> *I feel like it's the wrong question to ask, "how can we get black folks to trust these white researchers to come and, and, [laughter] and, and get their personal health information," right? I think the question really should be, "how can we find more black researchers? And how can we get more black doctors? And how can we set them up for success in areas where they're able to actually reach out with their community, right?" Like, the solution is not necessarily to place white doctors in black communities. That's an extra barrier. And as a researcher, I don't know why you would wanna do that.* [P4, 44:28]

Although participants shared an overall willingness to participate in a PrEP telehealth clinical trial led by Black investigators, they also shared that researchers generally should demonstrate care for the overall health and wellness of BSMM and not just recruit them for research.

Informed Consent: More Clarity Regarding Adverse Effects and Data Privacy

Despite information on how a consent form would detail the risks and privacy measures involved, every group mentioned that informed consent forms should provide more explicit, thorough, and understandable details. Participants were hesitant to believe that all known adverse effects regarding PrEP would be sufficiently outlined in a consent form for a study that was classified as an experiment. Initially, the extent to which PrEP efficacy was a part of the effectiveness of the clinical trial was not clear. Moreover, participants suggested that informed consent forms should explicitly outline intentions to not cause harm to participants. For example, one young man in a 25- to 34-year-old group said that the consent form should say, "Please. Thank you. I will not hurt you," which would be an example of the "care" for participants that BSMM suggested that researchers should demonstrate. Regarding data privacy, the older groups alluded to the need for consent forms to further describe "Who really gets access to my data?" They believed that their survey data and health history could easily be obtained by other researchers, clinicians, or pharmaceutical companies who were not a part of the study team.

Randomization As "Not Fair"

Participants across groups generally perceived that having a computer-generated process that randomized individuals into 1 of 2 groups was fair. However, 2 participants in an 18- to 24-year-old group mentioned that they would not participate in the clinical trial if they were randomized to a telehealth treatment group because they lived with their family members and would have privacy challenges. Some 25- to 34-year-old participants expressed preferences for the telehealth treatment arm because of the perception that it would be a better experience than the standard of care. Others did not want to be in an experimental telehealth arm because they feared that their data or laboratory information could be compromised because of the virtual format.

The 25- to 34-year-old participants suggested that randomization in this study is not fair to low-resourced BSMM who may not be able to fully participate in a long-term virtual study. They agreed with the man who thought that randomization to receive telehealth did not account for a participant's circumstance, specifically their ability to access a safe and private space to conduct a telehealth visit, to access a reliable internet connection, and to access technology.

> *I don't think it's fair--only because I think you have to think of the wholeness of the situation. If you're randomizing someone to the telehealth group, um, again, you have you have to see what type of resources that you have to give to them.* [P3, 36:24]

> *And they can privately be able to talk, you know? Because if they don't, it shuts out a lot of people in the community. So if we--they want it to be truly open to all people that if someone were to want to participate in the study and they are homeless, and they're assigned to telehealth. Then we can't guarantee them a private space or we can't guarantee that they have reliable internet connection. We can't guarantee that they have access to technology. Some people just do not feel comfortable sticking a needle in themselves, or whatever other, activities they need to further participate on their own. So I think that that's why the, the randomization, I think it's challenging, um, because there's so much more in consideration with this community.* [P3, 36:40]

Although only a few participants within the groups mentioned this, group members agreed that researchers could prevent a substantial subgroup of BSMM from participating in the telehealth clinical trial by not allowing them to choose their group assignment.

Wellness Check-Ins During Interim Visits

Participants suggested that the research team (either the staff or the principal investigator) should introduce themselves and conduct occasional wellness check-ins with participants regardless of race during interim study visits to build trust along the course of the study. They suggested that this type of communication would demonstrate "care" and investment in the overall health of BSMM. When asked to provide examples of how the research team should communicate with participants during the study, the 25- to 34-year-old participants said,

> *I believe that the researchers should make themselves available and present. I don't think they necessarily be at every single transaction throughout the course of the project. But they need to have, like, check-ins or midpoints and touchpoints 'cause that's the whole*

> *point. So, like, just even a follow-up like, "Hey, thanks for coming out, really appreciate what you're doing." Just, just something to let the people know that you actually care. Like, treat them like humans because if I only see you at the beginning of the orientation and at the end when it's all over, I'm gonna feel some kind of way because I feel like he didn't really care about me as a person.* [P4, 56:26]
>
> *Okay. Does that go for a researcher of any demographic or--?* [Facilitator, 56:58]
>
> *Any demographic.* [P4, 57:08]
>
> *Yeah. I feel like you should still reach out because I feel like you have to be more personal. If you didn't, then, I mean, I probably would never do a study with you again.* [P3, 57:24]

The 18- to 24-year-old participants said,

> *I guess email contact, maybe text messages, or maybe emails.* [S2, 36:57]
>
> *How often?* [Facilitator, 37:06]
>
> *Whatever the person is more comfortable with. I guess you just follow up, weekly or monthly or however you feel is necessary.* [S2: 37:06]
>
> *Yeah, I think being in touch and also letting them know that you are there for their safety and their health, that, yeah, they're going through a study, but ultimately, it's for their health and their well-being. Um, so being able to check on their health and their well-being, their mental health, you know, um, as well while they're doing the study just lets them know that, "Okay, you're committed to making sure that I'm given the support I need while helping you out." Because we're helping each other out, basically.* [S3, 37:22]

Free PrEP and Cash Incentives as Equitable

The 25- to 34-year-old participants explicitly mentioned that PrEP should be provided for free during the study as part of the incentives because of the low prevalence of adequate insurance coverage among BSMM. They suggested that it was unreasonable to expect participants in the study to pay for the medication along with the laboratory fees associated with PrEP care and that $50 cash incentives per visit would not be sufficient or equitable. When asked if providing larger cash incentives was coercive, participants across all the groups believed that cash incentives were more equitable for their time and participation in the study. The following is an example,

> *So what are y'all thoughts on monetary incentives for PrEP research? Like do you think people are being exploited when you give cash?* [Facilitator, 48:28]
>
> *No.* [P2, 48:43]
>
> *No.* [P4, 48:44]
>
> *No. I mean it's mutually beneficial to both parties, you know? I don't think that that's exploitative.* [P2, 48:46]
>
> *I mean people naturally want to get paid for the time, whether that be with money or some other form of incentive. So I don't think it's exploitative.* [PS2, 48:59]

Some participants also shared that participation in PrEP telehealth clinical trials was a way for some BSMM to obtain medical or financial support via clinical care and cash incentives.

> *Like I was saying before, a lot of people don't have the resources to have insurance. They can't afford insurance, some of them are only working part-time jobs. Or multiple part-time jobs that don't pay a lot in the first place. So I think it's kind of unfair to require that…maybe this research study is the only way I have to acquire this medication, acquire these resources because I can't afford the insurance.* [P4, 55:57]
>
> *Um, I think that if you're doing this on research, if you wanna include especially like LGBTQ people, gay people--* [P3, 57:05]
>
> *Young.* [P2, 57:18]
>
> *--then you should, you should, in that research program, find a way to provide, health care for free for at least a year. You know, some kind of--I don't know. I know that's a lot of work to get done, but I just feel like a lot of people in our community don't have access to health care. So if that's a requirement, then you're gonna be missing out on a lot of the people that, you know what I'm saying--* [P3, 57:19]
>
> *Right.* [P4, 57:28]

Ending Telehealth After Study as "Not Fair"

Participants across focus groups understood that PrEP telehealth resources would end with their participation and mentioned that it would be fair if the consent form explicitly stated that they would not be receiving the same resources after a specified amount of time. However, participants in the 25- to 34-year-old groups suggested that ending the convenience of PrEP telemedicine was not fair and would lower their trust and interest in future PrEP telehealth clinical trials. When asked whether not being able to have the telehealth treatment after the study ends is fair, one person said, "No. What the fuck? You can't get me used to something for a whole year and then just take it away." Others in the younger group said,

> *Not being able to have the services would definitely be, troubling if you just came to a study, and you were being helped, but now, all of a sudden, now your insurance is no longer being, paid for.* [P2, 44:50]
>
> *What do you think about being referred to another clinic for standard PrEP care once the study is over but you couldn't get telemedicine?* [Facilitator, 45:14]
>
> *At least you will be providing them, at least referring to somewhere where you can still get your medicine, even if it's not telemedicine.* [P2: 45:24]
>
> *I'm kinda thinking about my earlier comments about, like, you know, there's an office there? I think it would kinda be hard for people who have been--for a*

> *year--had the support they needed, and they're being cut off at the end of that year. So, I don't know. I think it's a tricky subject. I think that what they need to be sure is just to make sure they know at the end of this trial they may not have all of the support they used to have in that one year.* [P3, 45:34]

Participants viewed the study as a service provided to the community for their benefit, not necessarily an experimental treatment for a specific amount of time. They mentioned that collaborating with community-based organizations that could potentially continue similar services after the study increased their interest and trust in the study.

Discussion

This study explored culturally responsive study activities to build trust and improve the participation of BSMM in PrEP telehealth clinical trials. Overall, PrEP telehealth was an acceptable intervention strategy among BSMM. However, source of study funding, researchers' cultural congruence, intentions, and interactions, along with treatment assignment and ending telehealth impacted trust and study interest among BSMM. Medical and research mistrust persists in this population. The findings suggest that mistrust in PrEP telehealth clinical trials may persist because underlying issues regarding ethical clinical research conduct for minority groups have not been sufficiently addressed for BSMM. This study allows a reassessment of the traditionally acceptable domains of ethical research conduct to build trust and improve participation of BSMM in PrEP telehealth research.

The trust that BSMM have in PrEP telehealth clinical trials was assessed in part by their perception of how other and low-resourced BSMM may be treated or disregarded in the study. Concerns about the potential experiences of other in-group members is an important domain of history-based models of trust. Specifically, trust, according to this framework, is impacted by participants' own, vicarious, or anticipated experiences [21,30]. Cumulative negative interactions with society, family members, clinicians, and researchers that are experienced or expected could outweigh the perceived benefits of PrEP telehealth clinical trials and prevent study participation. These cumulative negative interactions could exacerbate mistrust in PrEP telehealth clinical trial research because trust has generally not been established in this group. More research is needed to understand how mistrust of PrEP telehealth clinical trials results from cumulative negative social and medical experiences because BSMM have historically been low-resourced and mistreated.

This study also revealed themes that established elements of care for BSMM throughout the PrEP telehealth clinical trial protocol and greater trust in a trial led by Black investigators. Other studies have found similar themes along the lines of establishing "care" to improve trust among BSMM [24,26,32]. Studies also showed that BSMM have greater trust in and less judgement from Black clinicians and researchers [24,26,38]. This finding is important because most clinical research teams and health care providers are not culturally congruent with this population [39,40]. Having PrEP telehealth clinical trials led by Black investigators could be an important understudied structural barrier to research participation among BSMM. More work is needed to increase the number of clinical trials led by Black investigators to assess the relative impact of this preference on PrEP uptake and study participation among BSMM.

Themes regarding the fairness of randomization to telehealth treatment groups and ending telehealth services suggest that BSMM assumed that the experimental treatment arm was inherently better than the standard of care and this impacted trust in the researchers and study. The assumption that one research group in a randomized controlled trial benefits more than the other undermines the presence of equipoise and raises questions regarding the ethical considerations of PrEP telehealth randomized controlled trial protocols [41]. Some researchers suggest that some randomized controlled trials are not necessarily investigated with equipoise and that most have directional hypotheses intended to demonstrate the effectiveness of one intervention over another [41,42]. Since recent clinical trials, including PrEP telehealth studies, intend to demonstrate some positive effect of the intervention over the standard of care, there are ethical considerations regarding equipoise that remain inadequately discussed. It is reasonable to think that BSMM would assume both that telehealth treatment is better than the care given to the control group and that the benefits of the treatment do not outweigh the effort involved in participation considering their history of marginalization and minimal resources. For traditionally low-resourced groups such as BSMM, PrEP clinical interventionists should reconsider study designs that have a group that receives "better treatment." Potentially, a single-arm pretest–posttest interventional study could be more appropriate when the assumption is that the treatment is "better" than the standard of care. Additionally, studies should request and obtain additional funding and budget to accommodate the culturally responsive activities that may be required to engage with this vulnerable population, such as providing technological devices and health care coverage for participants to sufficiently engage in the research. Traditional designs of randomized controlled trials might not be culturally responsive to the needs of BSMM and could perpetuate medical mistrust.

Importantly, we also found that information typically documented in an informed consent form (ie, minimal benefits, risks, privacy, random assignment, incentives, and end of study telehealth termination) was noted as unfair and insufficient. Guided by the Belmont Report [43], informed consent documents reflect the basic ethical principles of research conduct involving human subjects. The core tenets of the Belmont Report establish an imperative of informed consent detailing the nature of the study, benefits and risks, and randomization process and establishing participant comprehension prior to enrollment. However, data from the present study suggest that the ethical frameworks of justice and benefits within the Belmont Report [43] may require more thoughtful considerations and specificity for this subpopulation when PrEP telehealth clinical trials are conducted. Informed consent documents for PrEP telehealth clinical trial protocols may require tailoring to more adequately

identify what is meant by "comprehension" and better maximize "benefits" for BSMM participants.

The limitations of this study should be considered. This study did not quantify the prevalence of medical mistrust in this sample. Additionally, this convenience sample may have been biased toward a favorable attitude to PrEP telehealth generally because of participation in the virtual, synchronous focus group. Information from participants who were unwilling or unable to participate in the virtual focus group could have impacted the range of identified themes in this study. However, data from those participants are unavailable.

Overall, this study still provided important ethical and culturally responsive considerations for improving participation of BSMM in PrEP telehealth research. Given the salience of medical mistrust in the group, future studies should quantify the prevalence of these domains in the attitudes and willingness to participate in a PrEP telehealth clinical trial among BSMM. Future research should also assess the relative impact of implementing these findings on research participation in a PrEP telehealth clinical trial.

Acknowledgments

This research was supported in part by the National Institute of Mental Health Grant (no. K01MH11849 to principal investigator DTD II) and the Fordham University Research Ethics Training Institute Grant (no. R25DA031608 to principal investigator Celia Fisher). 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. Nelson LE. Awakening to an intersectional reality: ending the HIV epidemic in the USA starts with reducing inequities among Black MSM. J Urban Health 2020 Oct 06;97(5):589-591 [FREE Full text] [doi: 10.1007/s11524-020-00487-x] [Medline: 33025307]
2. Wong KYK, Stafylis C, Klausner JD. Telemedicine: a solution to disparities in human immunodeficiency virus prevention and pre-exposure prophylaxis uptake, and a framework to scalability and equity. Mhealth 2020 Apr;6:21-21 [FREE Full text] [doi: 10.21037/mhealth.2019.12.06] [Medline: 32270013]
3. Hess KL, Hu X, Lansky A, Mermin J, Hall HI. Lifetime risk of a diagnosis of HIV infection in the United States. Ann Epidemiol 2017 Apr;27(4):238-243 [FREE Full text] [doi: 10.1016/j.annepidem.2017.02.003] [Medline: 28325538]
4. HIV and African American gay and bisexual men | HIV by group | HIV/AIDS | CDC. Centers for Disease Control and Prevention. 2020 Feb 4. URL: https://www.cdc.gov/hiv/group/msm/bmsm.html [accessed 2020-10-21]
5. Anderson PL, Glidden DV, Liu A, Buchbinder S, Lama JR, Guanira JV, iPrEx Study Team. Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. Sci Transl Med 2012 Sep 12;4(151):151ra125 [FREE Full text] [doi: 10.1126/scitranslmed.3004006] [Medline: 22972843]
6. PrEP | HIV Basics | HIV/AIDS | CDC. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/hiv/basics/prep.html [accessed 2020-10-22]
7. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
8. Touger R, Wood BR. A review of telehealth innovations for HIV pre-exposure prophylaxis (PrEP). Curr HIV/AIDS Rep 2019 Feb 30;16(1):113-119. [doi: 10.1007/s11904-019-00430-z] [Medline: 30701404]
9. Hoots BE, Finlayson T, Nerlander L, Paz-Bailey G, National HIV Behavioral Surveillance Study Group. Willingness to take, use of, and indications for pre-exposure prophylaxis among men who have sex with men-20 US cities, 2014. Clin Infect Dis 2016 Sep 01;63(5):672-677 [FREE Full text] [doi: 10.1093/cid/ciw367] [Medline: 27282710]
10. Telehealth programs. Official web site of the U.S. Health Resources & Services Administration. URL: https://www.hrsa.gov/rural-health/telehealth [accessed 2020-10-22]
11. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017 Oct 19;377(16):1585-1592. [doi: 10.1056/NEJMsr1503323] [Medline: 29045204]
12. Refugio O, Kimble M, Silva C, Lykens J, Bannister C, Klausner J. Brief report: PrEPTECH: a telehealth-based initiation program for HIV pre-exposure prophylaxis in young men of color who have sex with men. A pilot study of feasibility. J Acquir Immune Defic Syndr 2019 Jan 01;80(1):40-45 [FREE Full text] [doi: 10.1097/QAI.0000000000001873] [Medline: 30272632]
13. Preexposure prophylaxis for the prevention of HIV infection in the United States-2017 Update: a clinical practice guidance. Centers for Disease Control and Prevention: US Public Health Service. Atlanta: Centers for Disease Control and Prevention; 2018 Mar. URL: https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-guidelines-2017.pdf [accessed 2020-05-11]

14. Hoth AB, Shafer C, Dillon DB, Mayer R, Walton G, Ohl ME. Iowa TelePrEP: A public-health-partnered telehealth model for human immunodeficiency virus preexposure prophylaxis delivery in a rural state. Sex Transm Dis 2019 Aug;46(8):507-512. [doi: 10.1097/OLQ.0000000000001017] [Medline: 31295217]
15. Quinn K, Dickson-Gomez J, Zarwell M, Pearson B, Lewis M. "A Gay Man and a Doctor are Just like, a Recipe for Destruction": how racism and homonegativity in healthcare settings influence PrEP uptake among young Black MSM. AIDS Behav 2019 Jul 18;23(7):1951-1963 [FREE Full text] [doi: 10.1007/s10461-018-2375-z] [Medline: 30565092]
16. Quinn K, Bowleg L, Dickson-Gomez J. "The fear of being Black plus the fear of being gay": The effects of intersectional stigma on PrEP use among young Black gay, bisexual, and other men who have sex with men. Soc Sci Med 2019 Jul;232:86-93 [FREE Full text] [doi: 10.1016/j.socscimed.2019.04.042] [Medline: 31075752]
17. Rolle C, Rosenberg ES, Siegler AJ, Sanchez TH, Luisi N, Weiss K, et al. Challenges in translating PrEP interest into uptake in an observational study of young Black MSM. J Acquir Immune Defic Syndr 2017 Nov 01;76(3):250-258 [FREE Full text] [doi: 10.1097/QAI.0000000000001497] [Medline: 28708811]
18. Cahill S, Taylor SW, Elsesser SA, Mena L, Hickson D, Mayer KH. Stigma, medical mistrust, and perceived racism may affect PrEP awareness and uptake in black compared to white gay and bisexual men in Jackson, Mississippi and Boston, Massachusetts. AIDS Care 2017 Nov;29(11):1351-1358 [FREE Full text] [doi: 10.1080/09540121.2017.1300633] [Medline: 28286983]
19. LaVeist T, Isaac L, Williams K. Mistrust of health care organizations is associated with underutilization of health services. Health Serv Res 2009 Dec;44(6):2093-2105 [FREE Full text] [doi: 10.1111/j.1475-6773.2009.01017.x] [Medline: 19732170]
20. Omodei MM, McLennan J. Conceptualizing and measuring global interpersonal mistrust-trust. J Soc Psychol 2000 Jun;140(3):279-294. [doi: 10.1080/00224540009600471] [Medline: 10902371]
21. Brandon DT, Isaac LA, LaVeist TA. The legacy of Tuskegee and trust in medical care: is Tuskegee responsible for race differences in mistrust of medical care? J Natl Med Assoc 2005 Jul;97(7):951-956. [Medline: 16080664]
22. Washington H. Medical Apartheid: The Dark History of Medical Experimentation on Black Americans from Colonial Times to the Present. New York City: Doubleday; Jan 09, 2007.
23. Arnett MJ, Thorpe RJ, Gaskin DJ, Bowie JV, LaVeist TA. Race, medical mistrust, and segregation in primary care as usual source of care: findings from the exploring health disparities in integrated communities study. J Urban Health 2016 Jun 18;93(3):456-467 [FREE Full text] [doi: 10.1007/s11524-016-0054-9] [Medline: 27193595]
24. White JJ, Dangerfield II DT, Grieb SM. Methodological considerations for conducting focus groups in HIV prevention research among Black men who have sex with men. Public Health Nurs 2019 May 13;36(3):439-445 [FREE Full text] [doi: 10.1111/phn.12592] [Medline: 30761620]
25. Arrington-Sanders R, Hailey-Fair K, Wirtz AL, Morgan A, Brooks D, Castillo M, et al. Role of structural marginalization, HIV stigma, and mistrust on HIV prevention and treatment among young Black Latinx men who have sex with men and transgender women: perspectives from youth service providers. AIDS Patient Care STDS 2020 Jan 01;34(1):7-15 [FREE Full text] [doi: 10.1089/apc.2019.0165] [Medline: 31944853]
26. Dangerfield II DT, Cooper J, Heidari O, Allen S, Winder TJ, Lucas GM. Nursing and health care preferences among opioid and stimulant using Black sexual minority men: an exploratory study. J Assoc Nurses AIDS Care 2020 Aug 17:e30-e39. [doi: 10.1097/JNC.0000000000000201] [Medline: 32809988]
27. Magnus M, Franks J, Griffith S, Arnold MP, Goodman K, Wheeler DP, HPTN 061 Study Group. Engaging, recruiting, and retaining black men who have sex with men in research studies: don't underestimate the importance of staffing--lessons learned from HPTN 061, the BROTHERS study. J Public Health Manag Pract 2014;20(6):E1-E9 [FREE Full text] [doi: 10.1097/PHH.0000000000000025] [Medline: 24406940]
28. Benoit E, Pass M, Randolph D, Murray D, Downing MJ. Reaching and engaging non-gay identified, non-disclosing Black men who have sex with both men and women. Cult Health Sex 2012 Oct 31;14(9):975-990 [FREE Full text] [doi: 10.1080/13691058.2012.709640] [Medline: 22937767]
29. Edmondson AC. Learning from failure in health care: frequent opportunities, pervasive barriers. Qual Saf Health Care 2004 Dec;13 Suppl 2:ii3-ii9 [FREE Full text] [doi: 10.1136/qhc.13.suppl_2.ii3] [Medline: 15576689]
30. Hammond W. Psychosocial correlates of medical mistrust among African American men. Am J Community Psychol 2010 Mar;45(1-2):87-106 [FREE Full text] [doi: 10.1007/s10464-009-9280-6] [Medline: 20077134]
31. Friedman MR, Sang JM, Bukowski LA, Chandler CJ, Egan JE, Eaton LA, et al. Prevalence and correlates of PrEP awareness and use among Black men who have sex with men and women (MSMW) in the United States. AIDS Behav 2019 Oct 28;23(10):2694-2705 [FREE Full text] [doi: 10.1007/s10461-019-02446-3] [Medline: 30820849]
32. Dangerfield II DT, Wylie C, Anderson JN. Conducting virtual, synchronous focus groups among Black sexual minority men: qualitative study. JMIR Public Health Surveill 2021 Feb 08;7(2):e22980 [FREE Full text] [doi: 10.2196/22980] [Medline: 33427671]
33. Ryan GW, Bernard HR. Techniques to identify themes. Field Methods 2016 Jul 24;15(1):85-109. [doi: 10.1177/1525822x02239569]
34. Ober AJ, Dangerfield II DT, Shoptaw S, Ryan G, Stucky B, Friedman SR. Using a "positive deviance" framework to discover adaptive risk reduction behaviors among high-risk HIV negative Black men who have sex with men. AIDS Behav 2018 May 13;22(5):1699-1712 [FREE Full text] [doi: 10.1007/s10461-017-1790-x] [Medline: 28501965]

35. Dangerfield II DT, Harawa NT, McWells C, Hilliard C, Bluthenthal RN. Exploring the preferences of a culturally congruent, peer-based HIV prevention intervention for black men who have sex with men. Sex Health 2018 Nov;15(5):424-430 [FREE Full text] [doi: 10.1071/SH18057] [Medline: 30185352]
36. Denzin KN, Lincoln YS. The SAGE Handbook of Qualitative Research (SAGE Handbooks). United States: SAGE, 2011; Apr 27, 2011:529-544.
37. Dangerfield II DT, Heidari O, Cooper J, Allen S, Lucas GM. Motivations for opioid and stimulant use among drug using black sexual minority men: A life course perspective. Drug Alcohol Depend 2020 Oct 01;215:108224 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108224] [Medline: 32777690]
38. Earl TR, Beach MC, Lombe M, Korthuis PT, Sharp V, Cohn J, et al. Race, relationships and trust in providers among Black patients with HIV/AIDS. Soc Work Res 2013 Sep 01;37(3):219-226 [FREE Full text] [doi: 10.1093/swr/svt017] [Medline: 24764690]
39. Sellers K, Leider J, Harper E, Castrucci BC, Bharthapudi K, Liss-Levinson R, et al. The Public Health Workforce Interests and Needs Survey: The First National Survey of State Health Agency Employees. J Public Health Manag Pract 2015;21 Suppl 6:S13-S27 [FREE Full text] [doi: 10.1097/PHH.0000000000000331] [Medline: 26422482]
40. Smiley RA, Lauer P, Bienemy C, Berg JG, Shireman E, Reneau KA, et al. The 2017 National Nursing Workforce Survey. J Nurs Regul 2018 Oct;9(3):S1-S88. [doi: 10.1016/s2155-8256(18)30131-5]
41. Cook C, Sheets C. Clinical equipoise and personal equipoise: two necessary ingredients for reducing bias in manual therapy trials. J Man Manip Ther 2011 Feb;19(1):55-57 [FREE Full text] [doi: 10.1179/106698111X12899036752014] [Medline: 22294855]
42. Gifford F. Uncertainty about clinical equipoise. Clinical equipoise and the uncertainty principles both require further scrutiny. BMJ 2001 Mar 31;322(7289):795 [FREE Full text] [Medline: 11282877]
43. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report. U.S. Department of Health & Human Services Office for Human Research Protections. 1979 Apr 18. URL: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html [accessed 2022-01-28]

Abbreviations

BSMM: black sexual minority men
PrEP: pre-exposure prophylaxis

Edited by A Kushniruk; submitted 15.03.21; peer-reviewed by M Almeida, B Wang; comments to author 02.10.21; revised version received 04.10.21; accepted 09.10.21; published 07.02.22.

Please cite as:
Dangerfield II DT, Wylie C
Identifying Ethical and Culturally Responsive Research Activities to Build Trust and Improve Participation of Black Sexual Minority Men in Pre-Exposure Prophylaxis Telehealth Clinical Trials: Qualitative Study
JMIR Hum Factors 2022;9(1):e28798
URL: https://humanfactors.jmir.org/2022/1/e28798
doi:10.2196/28798
PMID:35129448

Original Paper

Applications and User Perceptions of Smart Glasses in Emergency Medical Services: Semistructured Interview Study

Zhan Zhang[1], PhD; Karen Joy[1], BSc; Richard Harris[1], MSc; Mustafa Ozkaynak[2], PhD; Kathleen Adelgais[3], MD; Kevin Munjal[4], MD

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

[2]College of Nursing, University of Colorado, Aurora, CO, United States

[3]School of Medicine, University of Colorado, Aurora, CO, United States

[4]Department of Emergency Medicine, Mount Sinai Medical Center, New York, NY, United States

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: Smart glasses have been gaining momentum as a novel technology because of their advantages in enabling *hands-free* operation and *see-what-I-see* remote consultation. Researchers have primarily evaluated this technology in hospital settings; however, limited research has investigated its application in prehospital operations.

Objective: The aim of this study is to understand the potential of smart glasses to support the work practices of prehospital providers, such as emergency medical services (EMS) personnel.

Methods: We conducted semistructured interviews with 13 EMS providers recruited from 4 hospital-based EMS agencies in an urban area in the east coast region of the United States. The interview questions covered EMS workflow, challenges encountered, technology needs, and users' perceptions of smart glasses in supporting daily EMS work. During the interviews, we demonstrated a system prototype to elicit more accurate and comprehensive insights regarding smart glasses. Interviews were transcribed verbatim and analyzed using the open coding technique.

Results: We identified four potential application areas for smart glasses in EMS: enhancing teleconsultation between distributed prehospital and hospital providers, semiautomating patient data collection and documentation in real time, supporting decision-making and situation awareness, and augmenting quality assurance and training. Compared with the built-in touch pad, voice commands and hand gestures were indicated as the most preferred and suitable interaction mechanisms. EMS providers expressed positive attitudes toward using smart glasses during prehospital encounters. However, several potential barriers and user concerns need to be considered and addressed before implementing and deploying smart glasses in EMS practice. They are related to hardware limitations, human factors, reliability, workflow, interoperability, and privacy.

Conclusions: Smart glasses can be a suitable technological means for supporting EMS work. We conclude this paper by discussing several design considerations for realizing the full potential of this hands-free technology.

(JMIR Hum Factors 2022;9(1):e30883) doi:10.2196/30883

KEYWORDS

smart glasses; hands-free technologies; emergency medical services; user studies; mobile phone

Introduction

Background

Prehospital care is a high-risk, time-sensitive medical domain where first responders such as emergency medical services (EMS) providers provide urgent care to patients in the field and transport them to the nearest hospital or care facility. The primary goal of prehospital care is to stabilize patients by quickly addressing severe illnesses or life-threatening injuries. Perhaps prehospital care is among the most challenging medical domains in the provision of care to patients owing to various reasons, such as the broad range of clinical situations, limited resources and time, difficulties in accessing remote experts, and the highly dynamic situations and environmental conditions that providers encounter [1,2]. Owing to such challenges, technology support could be useful for EMS providers to facilitate decision-making and information management [3,4]. Despite some efforts, the prehospital environment remains one of the few medical settings with limited technology support [2]. In addition, previous work has primarily focused on developing and implementing systems on conventional handheld devices such as tablets or smartphones. For example, in an early study, Tollefsen et al [5] developed a menu-driven mobile app for EMS teams to document patient information, which was then uploaded to a central database for hospital care providers to instantaneously access and review. Another study designed and evaluated a smartphone app to facilitate care documentation in the field by enabling EMS providers to photograph the patient, record digital audio notes, and capture the view of the incident [6]. Despite their beneficial features, these handheld devices could cause problems in real-time use because (1) handheld devices are prone to interfere with manual tasks in a busy EMS environment [7-11] and (2) the physical handling of these devices could increase the chance of cross-contamination and patient infections [12]. As such, the need for novel technologies to support hands-busy EMS operations is evident [2,13].

Since being introduced to the public in 2011, smart glasses (a wearable technology in the shape of conventional glasses with a transparent screen and a video camera) have been gaining momentum because they can offer hands-free operation through novel interaction mechanisms such as voice control [14-18]. Google Glass has received the most exposure initially and stimulated the development of smart glasses by the industry. Compared with handheld devices, smart glasses enable constant information presentation and access in a hands-free manner and allow local workers to project first-person point-of-view to a remote viewer. Given these benefits, researchers have been exploring their potential in clinical and surgical environments [16], such as surgical telementoring [19-21], remote evaluation of patients with acute medical conditions [22], and disaster triage [23,24]. This body of literature demonstrated that smart glasses are potentially useful in supporting care management [16,17,25-28] and can enable secure Health Insurance Portability and Accountability Act (HIPAA)-compliant communications [29].

Research Gaps and Study Objective

Smart glasses can particularly benefit EMS as hands-free interaction can be a useful resource to handle situations with a lot of uncertainty. However, research on smart glasses in the prehospital environment is limited, with a few notable exceptions [13,23,24,30,31]. In addition, the previous work primarily focused on developing smart glass apps for certain EMS scenarios, such as disaster telemedicine triage [23,24], patient localization [30], and mobile vital sign monitoring [13]. To gain a comprehensive understanding of the application areas of smart glasses in EMS, the barriers and user concerns related to the use of smart glasses in practice, and how best to integrate this novel technology into the current EMS workflow, we conducted interviews to explore the potential and affordances of smart glasses in the out-of-hospital setting from the perspective of EMS providers to derive design implications for this novel technology. This study is part of a larger research effort that aims to iteratively design, develop, and evaluate smart glass technologies to support EMS operations in the field. In this work, we aim to answer the following three research questions through interviews with EMS providers: (1) How can smart glasses support EMS providers in overcoming challenges in the prehospital setting? (2) What interaction modality (eg, voice control, touch, and hand gestures) is most preferred and appropriate? (3) What types of concerns or potential barriers could impede the adoption and real-time use of this novel technology by EMS providers?

Methods

Study Design

We used a qualitative study approach (eg, interviews) [32,33] to gain an empirical and in-depth understanding of EMS providers' perceptions of and needs for adopting smart glasses in their daily work. This study approach has been successful in informing the design of complex sociotechnical systems [34]. The interview guide (Multimedia Appendix 1) was informed by previous work [35] and was developed in an iterative manner by the researchers. We also pilot-tested the interview guide with 2 experts (ie, EMS team leaders) to ensure the clarity, appropriateness, and relevance of the questions.

Participants

We conducted semistructured interviews with 13 EMS providers recruited from 4 hospital-based EMS agencies in an urban area in the east coast region of the United States. As shown in Table 1, a total of 85% (11/13) of them are paramedics, whereas the remaining 15% (2/13) are emergency medical technicians. Their years of experience ranged from 4 to 30 years, with 15% (2/13) of the participants being EMS directors. In addition, a few of them also serve other roles, such as EMS operation manager and quality assurance coordinator.

Table 1. Participant demographics (N=13).

ID	Sex	Occupation	Years of experience
P1	Male	Paramedic	28
P2	Male	Paramedic and EMS[a] educator	15
P3	Male	Paramedic and EMS director	25
P4	Male	Paramedic	18
P5	Male	Paramedic and quality assurance coordinator	30
P6	Male	Paramedic and EMS director	>30
P7	Female	Emergency medical technician	11
P8	Male	Paramedic	23
P9	Male	Paramedic	14
P10	Male	Emergency medical technician	4
P11	Male	Paramedic and EMS operation manager	21
P12	Male	Paramedic	11
P13	Male	Paramedic	7

[a]EMS: emergency medical services.

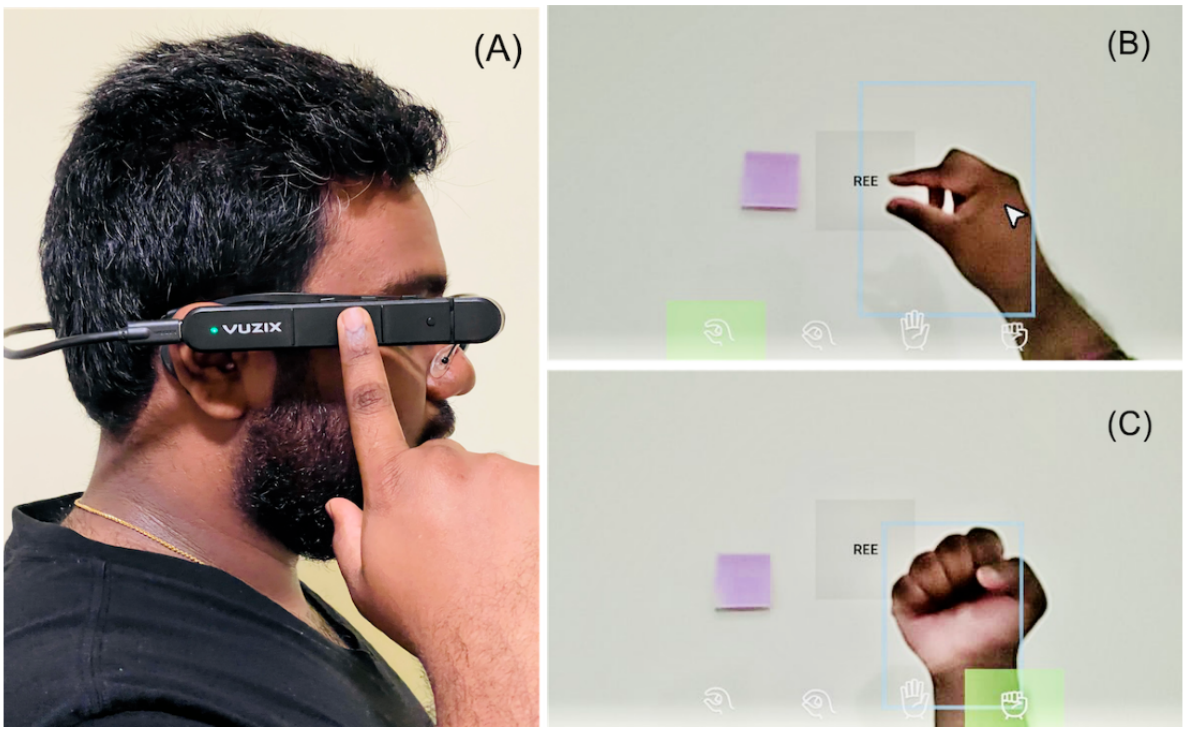
We included both emergency medical technicians and paramedics in our study because they represent the major types of EMS providers in the United States. Emergency medical technicians are trained to provide basic life support such as oxygen administration, wound treatment, and cardiopulmonary resuscitation. In contrast, the scope of practice and autonomy of paramedics are greater. Paramedics are allied health professionals with >1000 hours of training and provide advanced life support for patients, including advanced airway management, electrocardiogram interpretation, and medication administration. With both emergency medical technician and paramedic roles involved in our study, we were able to gain a holistic understanding of the use scenarios of smart glasses from different perspectives.

Data Collection

Owing to the COVID-19 pandemic, we conducted interviews via Zoom (Zoom Video Communications) following the best practices and experiences shared by other researchers who had to transition their user studies from in-person settings to web-based environments [36]. The interviews were conducted by two trained researchers (ZZ and KJ) and lasted for 45 to 90 minutes. Each interview was roughly divided into three sections: the first section consisted of general questions related to participant's demographics, work experience, and education or training background; the second section focused on EMS workflow from dispatch to patient hand off, artifacts and digital tools used in practice, and challenges encountered in their work (eg, documentation, care coordination, and communication with patients, dispatchers, and remote experts); the third section inquired about EMS professional's perceptions of using smart glasses in their daily work.

To help participants better understand this novel technology (eg, how it looks like and how it works) and elicit accurate and comprehensive insights, we used the Vuzix M400 [37] product to explain the hardware and software components of the device and possible interaction modalities through video and live demonstrations. In particular, we illustrated 3 modalities to interact with smart glasses. The first interaction modality was via the built-in touch pad and navigation buttons, which require physical touching and clicking on the device (Figure 1A). The second modality was voice commands created through the Vuzix software development kit. The third modality we demonstrated was hand gestures–based interaction, such as performing an open pinch to select the cursor (Figure 1B) and presenting an open hand and closing it to navigate back to the home page (Figure 1C). We implemented this interaction mechanism using the software development kit provided by CrunchFish [38], a Swedish technology company that develops gesture recognition software for mobile and wearable devices. We concluded the system demonstration by briefly discussing about its current application in other medical domains, such as in the operating room, wound care, and disaster triage.

Figure 1. Interaction modalities: (A) Use the built-in touch pad to navigate the user interface of smart glasses. (B) Perform an open pinch to select the cursor on the glass screen. (C) Present an open hand and close it to navigate back to the home screen.



Data Analysis

All interviews were audio-recorded and transcribed verbatim for analysis. Anonymized interview transcripts were analyzed by two researchers (ZZ and KJ) using an open coding technique [39]. We chose this coding technique because it could generate rich and detailed insights of participants' perspectives and opinions through iterative-inductive analysis of interview transcripts [40]. More specifically, the researchers first reviewed an initial set of transcripts independently (3/13, 23%). Then, the initial list of codes was generated and discussed among researchers to determine which codes to retain, merge, or remove. After the list of codes was set, we created a codebook to define each code to standardize the following coding process. The codebook was developed in an iterative manner until a consensus was reached. Then, the 2 researchers analyzed the remaining transcripts using the codebook. They met regularly to discuss and compare their codes for each interview transcript. The disagreements were resolved through discussion. Then, following a thematic analysis approach [41,42], the codes were grouped into themes using affinity diagrams [43], a common approach for creating connections or finding patterns in qualitative data. This step allowed the researchers to identify overarching themes describing EMS providers' opinions about the application areas of smart glasses in EMS operation, preferred interaction modalities, and potential barriers to adopting this novel technology. We have discussed about these major themes in the following section.

Ethical Consideration

This study was approved by the Pace University institutional review board (1515261). All participants provided their consent to participate in the study and be audio-recorded.

Results

Application Areas of Smart Glasses in EMS Work

Overview

The EMS participants identified a set of potential application areas where smart glasses could facilitate their work. We categorized them into four areas: teleconsultation, documentation, decision support, and quality assurance and training. We describe each use scenario in detail in the following sections.

Teleconsultation

The most prominent application area of smart glasses raised by EMS providers was teleconsultation. The reason is that EMS providers sometimes need to talk to a remote expert (eg, emergency department [ED] physician) for consultation, such as getting permission or advice for medication administration and collectively understanding the patient's status to decide the next steps. Currently, they rely solely on traditional communication mechanisms, such as radio or phone, to share and discuss information. However, these mechanisms have their intrinsic limitations, posing challenges for efficient communication between distributed EMS and ED teams. That is, in the field, EMS providers need to describe with words the situation they face; on the other side, ED physicians at the hospital often have difficulties in understanding what is precisely happening in the field potentially owing to ambient noise or disruptions in connectivity. As 1 participant explained:

> *Physicians maybe not hearing details correctly because it is over the radio. There's always going to be lag or miscommunication when you are using radios to relay information.* [P8]

As such, our participants expressed the emerging need for visual-based technologies to support their communication with the hospital. They believed that smart glasses could serve as an unobtrusive technological means to improve the communication and care coordination between prehospital and hospital care providers, as 1 participant explained:

> *I think it would be a useful tool, especially in those situations where you are going to end up contacting a physician, and they can actually see the environment in which the patient is. They can see the patient, specifically stroke patients, so that the physician can actually see the patient and facial droop and actually look at the patient. So, for those situations, I think it would be very helpful...You can have a conversation with the physician. Might be helpful for the physician to see the patient and what is being done to the patient at the same time for their purpose of understanding and getting a better picture of what's going on. I think it will be very helpful.* [P1]

Documentation

During the interviews, several participants identified the potential use of smart glasses in documentation. For example, participants stated that they could use smart glasses to take pictures and videos, which can be saved in the electronic health record (EHR) system to document patient injuries and wounds. In addition, these context-rich data can be shared with ED physicians to help them understand the severity of patient injury:

> *If you were able to do like a real-time video recording for a trauma patient to document like the mechanism of injury, like for falls or for car accident, and to be able to show those to the clinicians at the hospital, the doctors would love that.* [P4]

Another specific use for documentation is allowing EMS providers to dictate to the smart glasses and have the smart glasses transcribe the dictation to text through voice recognition. This use case was seen as a potential facilitator of documentation in the field, which could save a significant amount of time and efforts that can be spent on patient care:

> *I think it will be helpful for actual data collection for timing and all that. I think it's an excellent tool because it will make life easier if they [EMS providers] could actually just dictate certain things and they can be automatically stored in the electronic medical record.* [P5]

The ability to scan medications was also deemed useful by several participants. This feature could allow EMS providers to scan the barcode of the medication given to the patient, and the detailed information of the medication (eg, name and dosage) is automatically saved to EHR. In addition, as some patients could take several medications for chronic disease management and EMS providers may not have sufficient time to capture all the details, the medication scanning feature enabled by smart glasses could also make the collection of patient's medical history much easier:

> *Sometimes people come to us with a bag of medication and they're like, these are my medications. So, if I could just turn on the glasses and scan them and I have all the medications there...I mean, it'll not only make our jobs a little easier, but also expedite our patient's transport to the hospital.* [P7]

Decision Support

Participants believed that smart glasses could become a powerful decision support tool. For example, embedding medical protocols in the app could help EMS providers to perform a range of complex medical activities, verify the steps of less frequent tasks, and ensure compliance with medical procedures. One participant explained:

> *I think it would be great to integrate a point of reference to it (smart glass). Like, you know, we have protocols. Sometimes not everybody is going to have the same types of calls. Some people rarely get aphylactic calls, so when they get it, they would be like "oh my god, how much do I give?"...If you are newer in the field or you don't know your protocol because you've never encountered this problem, it would be great just to pull up that reference.* [P7]

Other mentioned decision support opportunities included augmenting information searching in a hands-free manner (P10), facilitating the determination of medication or fluid dosage (P12), and constantly presenting vital signs information to enhance situation awareness (P13):

> *Have a way to find references to normal vital signs or to look up definitions or features of different medical conditions.* [P10]

> *I think where the smart glasses could be very, very helpful is, as like a second check before you administer narcotics...If I don't have to take out my cell phone to confirm weight, pounds to kilograms and like, do the conversions, if I could just do that to my smart glasses without touching anything, I think that would be like perfect.* [P12]

> *If you were working on a cardiac arrest and the monitors not facing you, you can use the smart glass to pull up the vital signs using this head-up display. It would be awesome.* [P13]

Quality Assurance and Training

Participants considered smart glasses as a useful tool to record either the entire patient care process or critical medical procedures (eg, how the patient was treated for a complication in the field). This use of smart glasses can potentially enhance the quality of care by urging EMS providers to be more compliant:

> *If things are being recorded, it might make sure that you follow a protocol very correctly. With a digital record of when and how you do things, you're a lot more likely to kind of do things closer to the textbook.* [P13]

Furthermore, the video recordings of patient care can be used for both quality assurance and training purposes:

> *It'll be very good for educational quality assurance. You can actually trim videos for different things that went wrong, and you could use it to instruct your staff.* [P5]

Another interesting application area mentioned by a participant was helping with litigation issues faced by EMS providers in their work:

> *A lot of times, if something happens, you can be accused. But if you can have a recording with the voice, it could be used to protect the crews from litigation.* [P5]

Preferred Interaction Modalities

Participants were asked to rank the 3 demonstrated interaction modalities from most preferred to least preferred. Our data show that 46% (6/13) of the participants ranked voice commands and hand gestures equally as the most favored interaction mechanism:

> *I'd probably be an even mix of hand gestures and voice commands if that was like being used practically in real life. I think the voice control would be the preferred way to control the device, but in a loud situation the voice can probably be a little bit clunky. It'd be easier to use hand gestures to accomplish the same thing.* [P13]

Among the rest of the participants (7/13, 54%), 57% (4/7) of them chose voice commands as the most preferred modality, whereas 29% (2/7) of them voted the hand gesture modality.

In contrast, touch pad was indicated as the least preferred interaction modality, with only 14% (1/7) of the participants choosing this mechanism over the others. The reason for not favoring touch pad was either because it occupied hands or owing to concerns about cross-contamination:

> *Touch I think it's definitely bad cause your hands are always disgusting. Plus, with COVID going on right now, you touch one surface, then you're touching something that's very close to your face [smart glasses], that is not safe.* [P11]

Despite the various views on these interaction modalities, a few participants highlighted the importance of having all the interaction modalities available so that they can choose which one to use depending on the situation:

> *You have to have redundancy, you have to have a backup. So, let's say, it's not responding to your voice, then you can do the touch or whatever.* [P6]

Perceptions, Needs, and Concerns of Using Smart Glasses in Practice

Overview

Overall, our participants had a very positive attitude toward the use of smart glasses during prehospital encounters. Almost all of them expressed the willingness to use this system given its potential benefits; for example, not only supporting their work but also enhancing patient-centered care:

> *That would save so much time for the patient...So, you can do anything you got to do with your hands and it [smart glasses] will expedite the patient's care, which leads to a better patient outcome. I would definitely use it.* [P7]

Despite the positive attitude, participants shared several concerns about deploying this novel technology in practice. We grouped the major concerns into six categories, including hardware limitations, human factors, reliability, workflow, interoperability, and privacy.

Hardware Limitations

Battery life of smart glasses was a major concern:

> *They need to be able to last a while, making sure that it's not like going to run out on me in the middle of a high acuity situation.* [P12]

Participants suggested installing a charging station inside the ambulance and preparing 1-2 backup batteries to ensure that the smart glasses can constantly run through a whole work shift; that is, 8 hours.

Durability was another common concern expressed by participants because the prehospital environment is messy and fast-paced, where any device can easily get lost or broken:

> *The first impression to me is if it could survive in the 911 system...I constantly have equipment that breaks and malfunctions. It is unfortunately that the city 911 system is very rugged, abusive and tough. So, you need to make sure that something is tough and durable then you might stand the chance.* [P3]

Finally, smart glasses require a high-bandwidth cellular network to establish video calls; however, some areas (eg, rural areas and subway stations) rarely have high-speed network access. As such, use of smart glasses for teleconsultation could be impacted sometimes. Similarly, participants were concerned about not being able to transfer the recorded data from smart glasses to other devices (eg, EHR system) in a timely manner—an issue that is often caused by connection failure between devices. Therefore, a few participants mentioned that smart glasses should have sufficient internal memory to allow *store-and-forward*, a common data transferring method in many telemedicine systems [44]:

> *If I'm in subway and there is no WIFI, and if I record videos or take pictures, or you know, telling it to dictate something, it should have enough internal memory to store everything recorded.* [P5]

Other hardware-related factors that might impact the adoption of smart glasses included device cost (eg, "my only concern is how much will this cost per unit?" [P1]), process of disinfecting the device (eg, "If it gets contaminated, how do you clean it?" [P5]), and safety issues (eg, "I'm definitely concerned about getting assaulted with it on." [P9]).

Human Factors

Our participants raised several issues with regard to human factors. For example, 15% (2/13) of the participants commented

that smart glasses could block a certain field of vision and, in turn, affect their patient care activities:

> *I think it could impede work if it's obstructing my view. I think that the most likely scenario where there would truly be an impedance to patient care is intubating a patient because I need to make sure that I have good vision of the patient's vocal cords.* [P12]

In addition, it is vital to ensure that the device is not intrusive because "after a certain amount of time, stuff on your head could get irritating or annoying" [P5].

Compatibility with users' own glasses or personal protective equipment was also frequently mentioned by our participants:

> *I like to wear goggles now, especially because of Covid-19. You need to make sure that the glasses have a good fit with the rest of your PPE, whether it's a goggle or a face mask.* [P12]

Reliability

The smart glass app must be reliable because EMS work is high-acuity and time-critical; any system malfunction could lead to increased stress and high cognitive workload on EMS providers and even adverse patient outcomes. As such, system reliability is one of the primary concerns expressed by multiple participants:

> *If there is a technological failure, what is our backup, what do we do?* [P4]

Workflow

As our participants had little experience with smart glasses (compared with mobile phones or tablets), they were not clear whether this novel technology can seamlessly fit into EMS workflow:

> *Just like with every technology, just making sure that it is seamless and actually works. All of our technologies make sense in theory, but the application can be a little bit difficult.* [P12]

Another participant shared the same opinion and further commented that smart glasses might add more workload, such as the need to check the accuracy of recorded data when using it for documentation:

> *Obviously they [EMS providers] have to make sure that the information was recorded correctly. Is that another component that's going to add time?* [P3]

Interoperability

Interoperability was also mentioned by the participants. For example, smart glasses should be integrated with the EHR system to realize the documentation use. Similarly, timely and constant data exchange (eg, electrocardiography, blood pressure, and oxygen saturation) between smart glasses and the vital signs monitor is critical for implementing the decision support feature. As such, the interoperability between smart glasses and other medical devices is essential:

> *I guess compatibility to different devices is very important. You know, trying to integrate something as simple as a monitor to any sort of technology is a bit of a hurdle because they don't play nicely with each other.* [P2]

Privacy

Not surprisingly, data privacy was one of the most prominent concerns shared by almost all participants. Given that smart glasses would transfer, retrieve, and even store sensitive patient data, it is imperative to ensure compliance with the HIPAA regulations:

> *The number one concern would be patient privacy. That would have to be worked out. How is the data getting stored? How is the data getting processed?* [P4]

They also have concerns that patients, especially pediatric patients, might feel uncomfortable or nervous while seeing them wear a pair of smart glasses, as this device is rarely seen in daily life. Sometimes, patients may not even want to be recorded, so participants suggested that new regulations or rules regarding the digital recording of patients should be established before deploying the system in the field:

> *You always might have some patients who may not want to be recorded unless that's the policy...you may also need to get patient's consent and it becomes the legal issue.* [P1]

Finally, as smart glasses can capture the conversations, actions, and patient care process in videos, several participants expressed concerns about their personal liability and felt that they would be working under observation:

> *Maybe stuff you don't want recorded gets recorded and then a supervisor uses that against you. If they're trying to look for something, you know. It goes back to like the same things with police body cams, where, you know, stuff gets recorded without their knowledge. Sometimes it doesn't turn on when it's supposed to. So those would be the concerns I had and then who has access to see and hear what's recorded.* [P8]

Discussion

Principal Findings

In this study, we conducted semistructured interviews with prehospital care providers to understand the potential and affordances of smart glasses in EMS. We identified several potential application areas for smart glasses to support EMS work in the field, including (1) enhancing communication and consultation between distributed prehospital and hospital providers, (2) supporting patient data collection and documentation in a hands-free manner, (3) supporting decision-making and situation awareness, and (4) augmenting quality assurance and training. In the following section, situated in previous work, we discuss the feasibility of these potential applications and design considerations for realizing them. Major design considerations for the 4 identified application areas are summarized in Textbox 1.

Textbox 1. Summary of application areas and design considerations for applying smart glasses in emergency medical services.

Teleconsultation

- Smart glasses should be designed to augment rather than replace current communication tools.
- Advanced mounting techniques are needed to make sure smart glasses sit steadily in front of the user's eyes.

Documentation

- Novel techniques are needed to enable high performance of automatic speech recognition feature of smart glasses.
- More tests are needed to examine the usability and affordances of smart glasses in transcribing medical procedures.

Decision support

- Smart glass–based decision support interventions (eg, checklist) need to be designed such that they are dynamic and flexible enough to adapt to different patient scenarios.
- Artificial intelligence, computer vision, and smart glasses should be combined to automatically detect a patient's signs and symptoms.

Quality assurance and training

- Patient data security and confidentiality must be maintained in accordance with Health Insurance Portability and Accountability Act regulations.
- Rules and policies need to be enacted to guide when video recording is allowed and who has the permission to watch the videos.

Communication and care coordination between prehospital and hospital teams are essential for safe, timely, and effective patient care. For example, the treatment of a pediatric patient with traumatic brain injury with a rapidly changing state of consciousness often requires a considerable level of knowledge and skills that EMS providers may not have. In this case, EMS providers may need to consult with a more experienced ED physician for advice (eg, what medications to administer or how to perform treatments that are critical to save the patient's lives during ambulance transport) [45,46]. Furthermore, smooth communication can enable efficient joint decision-making between EMS and ED care providers with regard to the treatment plan, likely diagnoses, and appropriate destination of care [6,47-49]. Despite its critical role, this process remains ineffective [50-53]. This challenge is owing in part to the limitations of current communication mechanisms (eg, radio) because they limit multisensorial interaction—an important mechanism for ensuring smooth work and cooperation among collaborators [54]—between distributed care providers. Our study revealed that smart glasses were perceived to be a useful tool for EMS providers to connect with remote experts because they fulfill the need of visual supports through a *see-what-I-see* video. In fact, previous work has revealed the usefulness and feasibility of smart glasses in establishing remote expert support, such as in surgical telementoring [19-21], remote evaluation of patients experiencing acute stroke [22], and disaster telemedicine triage [23,24,55]. For example, a recent study [55] indicated that using smart glasses led to increased quality of triage during mass casualty incidents (MCIs). In addition, EMS providers reported satisfactory usability and good acceptance of the smart glass technology. However, there are a few considerations for deploying smart glasses in the out-of-hospital setting for use by EMS providers. For example, because smart glasses require a high-bandwidth cellular network for video calls, which is rare in some places (eg, rural areas and subway stations), smart glasses should only augment rather than replace current communication tools (eg, radio or cellular phone). However, with the rapid development of 5G technology and the proposition of building a dedicated broadband network for first responders (eg, FirstNet [56]), this limitation might be addressed in the near future. Another design consideration is regarding a common problem of using smart glasses for teleconsultation—difference in line of sight between distributed collaborators; that is, the remote expert could not always see what exactly the smart glass wearer's local eyes were fixed on [57]. Therefore, more advanced mounting techniques are needed to ensure that smart glasses sit steadily in front of the user's eyes even during excessive physical activities.

Collecting and documenting patient data in the field is a challenging and time-consuming task, which demands a significant portion of EMS professionals' cognitive attention, thereby reducing their time spent on patient care [58]. Despite the increased adoption of EHR systems by EMS agencies, the real-time use of the EHR systems has faced many challenges. For example, as these systems are implemented on handheld devices such as tablets, EMS providers may not be able to use such devices in real time given the dynamic and hands-busy nature of prehospital care [10,11]. In addition, the use of handheld devices could increase the chance of cross-contamination [12]. Compared with handheld devices, smart glasses offer advantages such as hands-free operation, which has the potential to support real-time patient data collection and documentation. However, to date, very few studies have focused on supporting clinical documentation using smart glasses [28,59,60]. For example, a previous work [60] reported the design and evaluation of a smart glass app for chronic wound photography, which supported wound care nurses in documentation by enabling capture, tagging, and transfer of images to a patient's EHR in a hands-free manner. In another study [59], researchers tested the patient's acceptance and perception of their physician wearing smart glasses to connect with a remote scribe nurse who took notes during a clinical visit. These studies demonstrated the usefulness of smart glasses in supporting timely clinical documentation. However, they were conducted in settings where the working stress and noise level

are considerably lower than the prehospital domain, which is often characterized as a noisy and messy environment that could affect the effective use of the voice recognition feature of smart glasses. In recent years, novel techniques have been developed to address this issue, such as a sensing and signal processing solution that enables high performance of automatic speech recognition of smart glasses [61]. To overcome challenges in realizing the documentation use for EMS, future research is needed to systematically test the usability and affordances of smart glasses in transcribing medical procedures while EMS providers perform them in noisy, dynamic, and fast-paced environments.

Similarly, very limited research has been conducted to examine the use of smart glasses for decision-making support. The most common feature reported in previous work is the presentation of a checklist or medical protocols on the glass screen. For example, in a recent study [62], researchers compared a smart glass–based checklist with conventional methods (ie, memory or poster) during surgical cases and found that smart glass–based checklists increased the completion rate to 100% and reduced the time required to execute the checklist and prepare the equipment. Another study [63] implemented triage algorithms on the smart glass platform to support the triage process during MCIs and reported that most EMS participants found the triage app to be useful or partially useful. Similarly, Follmann et al [55] found that smart glass can improve triage results during an MCI by showing the triage algorithms and by receiving support from a physician. These previous studies, despite not being extensive, illustrated that EMS teams could benefit greatly from smart glass checklists [3]. Given the unique characteristics of the EMS environment (eg, unpredictable clinical scenarios), smart glass–based checklists need to be designed such that they are dynamic and flexible enough to adapt to different patient scenarios, including less frequent but critical tasks [64,65]. In addition to the checklist application, future research can also explore combining artificial intelligence, computer vision, and smart glasses to automatically detect a patient's signs and symptoms and recommend treatment options accordingly. For example, the artificial intelligence–powered smart glasses can help EMS providers to identify early signs of critical illnesses (eg, stroke) or hard-to-detect mechanisms of injury (eg, child abuse).

The use of smart glasses for quality assurance has received little attention so far, but it could become a new application area of smart glasses for not only EMS but also other medical domains. This use is realized mainly through the video recording feature. However, the challenging part is related to privacy issues. First, patient data security and confidentiality must be maintained in accordance with HIPAA regulations and other local, federal, or organizational policies. Second, patients and their surrounding environment (including bystanders) can be captured by smart glasses without their knowledge; that is, when transmitting videos from the field to the hospital for consultation. Therefore, EMS providers may need to obtain verbal or even written consent from the patient before using video recordings or calls. The smart glasses should also be designed to protect bystanders' identities and privacy, such as automatically blurring their faces or recognizing their hand gestures for signaling consent (opt in) or disapproval (opt out) [66]. In addition, the glasses should clearly indicate when they are capturing videos to increase the awareness of bystanders; that is, through a light emitting diode strip [31,67]. Finally, as our participants explained, using smart glasses to capture EMS providers' conversations, actions, and patient care processes in videos could easily trigger their privacy and liability concerns. Given these privacy issues and considerations that are entailed in using smart glasses to record videos, organizational and national rules and regulations should be in place to provide guidance as to when video recording is allowed and who has the permission to watch the videos.

Regardless of the application areas, users should be able to interact with the smart glasses in an intuitive manner without disrupting their work practice. Our data shows that voice commands and hand gestures are preferred over touch pad because of their *hands-free* advantage. Despite the benefits, noisy environment can pose challenges in using voice commands, whereas fast-moving ambulances can affect the use of hand gestures. Therefore, as our participants stated, it is critical for them to have the option to use a mix of interaction mechanisms at any time while using the smart glass app.

Our study also revealed some other considerations that need full attention before deploying the smart glass technology in EMS. For example, as previous work has pointed out [35], it is important to ensure that the device's battery can last long enough for care providers to use it throughout a shift. In addition, interoperability was cited as a critical consideration for the successful implementation of smart glasses in the field. That is, the smart glass device should be integrated with existing systems (eg, vital signs monitor and EHR system) to allow seamless data exchange. Finally, the medicolegal aspects of smart glasses could impact the real use of this novel technology. For example, when using smart glasses as a telemedicine tool, the medicolegal obligations are placed on both distant and local emergency care providers. In addition, the patient must be informed about the nature, purpose, and use of the smart glass device and what benefits this technology can offer to them [68]. Any potential breach in patient's privacy and confidentiality must be addressed to enhance patient-centered care.

Limitations

This study has several limitations that need to be noted. First, the interviews were not conducted in person. Although we gave video and live demonstrations of the smart glass device, participants did not get an opportunity to use it. This limitation could have impacted their views on this technology. Second, participants were recruited from hospital-based EMS agencies in an urban area of the east coast region in the United States. Therefore, the user perceptions were based on how they operate locally, which may be different from other places (eg, rural areas and other regions of United States) and other types of EMS providers (eg, fire department–based or volunteer-based), let alone other countries. Therefore, the results may not be generalizable to all types of EMS agencies worldwide. More work in other regions and countries is needed to supplement the findings of this study. Third, user opinions and needs were collected only through interviews. We neither asked the participants to use the device nor tested the effectiveness and

usefulness of this technology under different case scenarios. This may have impacted the participants' views on smart glass. Additional studies, such as participatory design workshop and usability evaluation, will be conducted in the future to elicit additional design insights about smart glasses for EMS. In addition, it is critical to conduct simulated scenarios to test the efficiency and effectiveness of smart glasses for different application areas. Finally, our participants were mostly male. Female participants may have different opinions and preferences. In our future work, we will include as many female EMS providers as possible.

Conclusions

In this study, we conducted semistructured interviews with EMS providers to learn their opinions, needs, and concerns regarding the use of smart glasses in their daily work. Our results identified four potential application areas in which smart glasses can play an essential role, including enhancing teleconsultation between distributed prehospital and hospital providers; semiautomating patient data collection and documentation in real time; aiding decision-making and situation awareness; and finally, augmenting quality assurance and training. We also found that voice commands and hand gestures were preferred over the built-in touch pad for system navigation. Although EMS providers consider smart glasses as a suitable technological means for prehospital work, several issues and user concerns, such as hardware limitations, human factors, reliability, workflow, interoperability, and privacy, need to be thoroughly addressed to ensure its successful uptake and implementation. Finally, we identified several design considerations for realizing the applications of smart glasses in EMS.

Acknowledgments

This study was supported by the National Science Foundation award 1948292 and the Agency for Healthcare Research and Quality award 1R21HS028104-01A1.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 19 KB - humanfactors_v9i1e30883_app1.docx]

References

1. Sarcevic A, Zhang Z, Kusunoki D. Decision making tasks in time-critical medical settings. In: Proceedings of the 17th ACM International Conference on Supporting Group Work. 2012 Presented at: GROUP '12; October 27-31, 2012; Sanibel island, FL p. 99-102. [doi: 10.1145/2389176.2389191]
2. Zhang Z, Sarcevic A, Bossen C. Constructing common information spaces across distributed emergency medical teams. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR p. 934-947. [doi: 10.1145/2998181.2998328]
3. Ozkaynak M, Dolen C, Dollin Y, Rappaport K, Adelgais K. Simulating teamwork for better decision making in pediatric emergency medical services. AMIA Annu Symp Proc 2021;2020:993-1002 [FREE Full text] [Medline: 33936475]
4. Zhang Z, Sarcevic A, Burd RS. Supporting information use and retention of pre-hospital information during trauma resuscitation: a qualitative study of pre-hospital communications and information needs. AMIA Annu Symp Proc 2013;2013:1579-1588 [FREE Full text] [Medline: 24551428]
5. Tollefsen WW, Gaynor M, Pepe M, Myung D, Welsh M, Moulton S. iRevive: a pre-hospital database system for emergency medical services. Int J Healthc Technol Manag 2005;6(4-6):454-469. [doi: 10.1504/ijhtm.2005.007008]
6. Schooley B, Abed Y, Murad A, Horan TA, Roberts J. Design and field test of an mHealth system for emergency medical services. Health Technol 2013;3(4):327-340. [doi: 10.1007/s12553-013-0064-5]
7. Rogers H, Madathil KC, Agnisarman S, Narasimha S, Ashok A, Nair A, et al. A systematic review of the implementation challenges of telemedicine systems in ambulances. Telemed J E Health 2017;23(9):707-717. [doi: 10.1089/tmj.2016.0248] [Medline: 28294704]
8. Cho SJ, Kwon IH, Jeong J. Application of telemedicine system to prehospital medical control. Healthc Inform Res 2015;21(3):196-200 [FREE Full text] [doi: 10.4258/hir.2015.21.3.196] [Medline: 26279957]
9. Chapman Smith SN, Brown PC, Waits KH, Wong JS, Bhatti MS, Toqeer Q, et al. Development and evaluation of a user-centered mobile telestroke platform. Telemed J E Health 2019;25(7):638-648 [FREE Full text] [doi: 10.1089/tmj.2018.0044] [Medline: 30207927]
10. Jagannath S, Sarcevic A, Multak N, Myers S. Understanding paper-based documentation practices in medical resuscitations to inform the design of electronic documentation tools. Pediatr Emerg Care 2021;37(8):e436-e442. [doi: 10.1097/PEC.0000000000001676] [Medline: 30586038]

11. Sarcevic A, Ferraro N. On the use of electronic documentation systems in fast-paced, time-critical medical settings. Interact Comput 2017;29(2):203-219. [doi: 10.1093/iwc/iww021]
12. Momen KS. Identifying nursing activities to estimate the risk of cross-contamination. University of Toronto. 2013. URL: https://tspace.library.utoronto.ca/handle/1807/34914 [accessed 2022-02-24]
13. Schaer R, Melly T, Muller H, Widmer A. Using smart glasses in medical emergency situations, a qualitative pilot study. In: 2016 IEEE Wireless Health. 2016 Presented at: WH '16; October 25-27, 2016; Bethesda, MD p. 1-5. [doi: 10.1109/wh.2016.7764556]
14. Noma H, Ohmura A, Kuwahara N, Kogure K. Wearable sensors for auto-event-recording on medical nursing-user study of ergonomic design. In: 8th International Symposium on Wearable Computers. 2004 Presented at: ISWC '04; October 31-November 3, 2004; Arlington, VA p. 8-15. [doi: 10.1109/iswc.2004.48]
15. Jonas S, Hannig A, Spreckelsen C, Deserno TM. Wearable technology as a booster of clinical care. In: Proceedings of the SPIE 9039, Medical Imaging 2014: PACS and Imaging Informatics. 2014 Presented at: SPIE '14; February 15-20, 2014; San Diego, CA p. 90390F. [doi: 10.1117/12.2042986]
16. Mitrasinovic S, Camacho E, Trivedi N, Logan J, Campbell C, Zilinyi R, et al. Clinical and surgical applications of smart glasses. Technol Health Care 2015;23(4):381-401. [doi: 10.3233/THC-150910] [Medline: 26409906]
17. Klein GO, Singh K, von Heideken J. Smart glasses--a new tool in medicine. Stud Health Technol Inform 2015;216:901. [Medline: 26262203]
18. Carlsson V, Klug T, Ziegert T, Zinnen A. Wearable computers in clinical ward rounds. In: 3rd International Forum on Applied Wearable Computing. 2006 Presented at: IFAWC '06; March 15-16, 2006; Bremen, Germany.
19. Christensen NH, Hjermitslev OG, Falk F, Madsen MB, Østergaard FH, Kibsgaard M, et al. Depth cues in augmented reality for training of robot-assisted minimally invasive surgery. In: Proceedings of the 21st International Academic Mindtrek Conference. 2017 Presented at: AcademicMindtrek '17; September 20-21, 2017; Tampere, Finland p. 120-126. [doi: 10.1145/3131085.3131123]
20. Lee JS, Tsai CT, Pen CH, Lu HC. A real time collaboration system for teleradiology consultation. Int J Med Inform 2003;72(1-3):73-79. [doi: 10.1016/s1386-5056(03)00130-8] [Medline: 14644308]
21. Weibel N, Gasques D, Johnson J, Sharkey T, Xu ZR, Zhang X, et al. Artemis: mixed-reality environment for immersive surgical telementoring. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI EA '20; April 25-30, 2020; Honolulu, HI p. 1-4. [doi: 10.1145/3334480.3383169]
22. Noorian AR, Bahr Hosseini M, Avila G, Gerardi R, Andrle AF, Su M, et al. Use of wearable technology in remote evaluation of acute stroke patients: feasibility and reliability of a Google glass-based device. J Stroke Cerebrovasc Dis 2019;28(10):104258. [doi: 10.1016/j.jstrokecerebrovasdis.2019.06.016] [Medline: 31296476]
23. Broach J, Hart A, Griswold M, Lai J, Boyer EW, Skolnik AB, et al. Usability and reliability of smart glasses for secondary triage during mass casualty incidents. Proc Annu Hawaii Int Conf Syst Sci 2018;2018:1416-1422 [FREE Full text] [Medline: 29398976]
24. Cicero MX, Walsh B, Solad Y, Whitfill T, Paesano G, Kim K, et al. Do you see what I see? Insights from using Google glass for disaster telemedicine triage. Prehosp Disaster Med 2015;30(1):4-8. [doi: 10.1017/S1049023X1400140X] [Medline: 25571779]
25. Aungst TD, Lewis TL. Potential uses of wearable technology in medicine: lessons learnt from Google glass. Int J Clin Pract 2015;69(10):1179-1183. [doi: 10.1111/ijcp.12688] [Medline: 26412708]
26. Wu TS, Dameff CJ, Tully JL. Ultrasound-guided central venous access using Google glass. J Emerg Med 2014;47(6):668-675. [doi: 10.1016/j.jemermed.2014.07.045] [Medline: 25281180]
27. Yu J, Ferniany W, Guthrie B, Parekh SG, Ponce B. Lessons learned from Google glass: telemedical spark or unfulfilled promise? Surg Innov 2016;23(2):156-165. [doi: 10.1177/1553350615597085] [Medline: 26224576]
28. Klinker K, Wiesche M, Krcmar H. Digital transformation in health care: augmented reality for hands-free service innovation. Inf Syst Front 2020;22(6):1419-1431. [doi: 10.1007/s10796-019-09937-7]
29. Chai PR, Wu RY, Ranney ML, Porter PS, Babu KM, Boyer EW. The virtual toxicology service: wearable head-mounted devices for medical toxicology. J Med Toxicol 2014;10(4):382-387 [FREE Full text] [doi: 10.1007/s13181-014-0420-5] [Medline: 25135308]
30. Reuter C, Ludwig T, Mischur P. RescueGlass: collaborative applications involving head-mounted displays for red cross rescue dog units. Comput Supported Coop Work 2019;28(1-2):209-246. [doi: 10.1007/s10606-018-9339-8]
31. Schlosser P, Matthews B, Salisbury I, Sanderson P, Hayes S. Head-worn displays for emergency medical services staff: properties of prehospital work, use cases, and design considerations. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-14. [doi: 10.1145/3411764.3445614]
32. Fetterman DM. Ethnography: step-by-step. Thousand Oaks, CA: Sage Publications; 2019.
33. Spradley JP. The ethnographic interview. Long Grove, IL: Waveland Press; 2016:1478633212.
34. Randall D, Harper R, Rouncefield M. Fieldwork for design: theory and practice. Berlin, Germany: Springer Science & Business Media; 2007.

35. Romare C, Hass U, Skär L. Healthcare professionals' views of smart glasses in intensive care: a qualitative study. Intensive Crit Care Nurs 2018;45:66-71 [FREE Full text] [doi: 10.1016/j.iccn.2017.11.006] [Medline: 29291922]
36. Mastrianni A, Kulp L, Sarcevic A. Transitioning to remote user-centered design activities in the emergency medical field during a pandemic. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI EA '21; May 8-13, 2001; Yokohama, Japan p. 1-8. [doi: 10.1145/3411763.3443444]
37. Vuzix M400. Vuzix. URL: https://www.vuzix.com/products/m400-smart-glasses [accessed 2021-11-20]
38. Crunchfish. URL: https://www.crunchfish.com [accessed 2021-11-20]
39. Holton JA. The coding process and its challenges. In: Bryant A, Charmaz K, editors. The Sage handbook of grounded theory. Thousand Oaks, CA: Sage Publications; 2007:265-289.
40. O'reilly K. Ethnographic methods. 2nd edition. Milton Park, UK: Routledge; 2012.
41. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006;5(1):80-92. [doi: 10.1177/160940690600500107]
42. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1609406917733847. [doi: 10.1177/1609406917733847]
43. Hartson R, Pyla PA. The UX book: process and guidelines for ensuring a quality user experience. San Diego, CA: Morgan Kauffman; 2012.
44. Tulu B, Chatterjee S, Maheshwari M. Telemedicine taxonomy: a classification tool. Telemed J E Health 2007;13(3):349-358. [doi: 10.1089/tmj.2006.0055] [Medline: 17603838]
45. Emberson J, Lees KR, Lyden P, Blackwell L, Albers G, Bluhmki E, Stroke Thrombolysis Trialists' Collaborative Group. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. Lancet 2014;384(9958):1929-1935 [FREE Full text] [doi: 10.1016/S0140-6736(14)60584-5] [Medline: 25106063]
46. Audebert HJ, Schwamm L. Telestroke: scientific results. Cerebrovasc Dis 2009;27 Suppl 4:15-20. [doi: 10.1159/000213054] [Medline: 19546537]
47. Aase K, Soeyland E, Hansen BS. A standardized patient handover process: perceptions and functioning. Saf Sci Monit 2011;15(2):1-9.
48. Sarcevic A, Burd RS. Information handover in time-critical work. In: Proceedings of the ACM 2009 International Conference on Supporting Group Work. 2009 Presented at: GROUP '09; May 10-13, 2009; Sanibel island, FL p. 301-310. [doi: 10.1145/1531674.1531720]
49. Bergrath S, Rörtgen D, Rossaint R, Beckers SK, Fischermann H, Brokmann JC, et al. Technical and organisational feasibility of a multifunctional telemedicine system in an emergency medical service - an observational study. J Telemed Telecare 2011;17(7):371-377. [doi: 10.1258/jtt.2011.110203] [Medline: 21933897]
50. Schuster M, Pints M, Fiege M. Duration of mission time in prehospital emergency medicine: effects of emergency severity and physicians level of education. Emerg Med J 2010;27(5):398-403. [doi: 10.1136/emj.2009.074211] [Medline: 20442179]
51. McDonald KM, Schultz E, Albin L, Pineda N, Lonhart J, Sundaram V, et al. Care coordination measures atlas update. Agency for Healthcare Research and Quality (AHRQ). 2014. URL: https://www.ahrq.gov/ncepcr/care/coordination/atlas.html [accessed 2022-02-24]
52. Schultz EM, Pineda N, Lonhart J, Davies SM, McDonald KM. A systematic review of the care coordination measurement landscape. BMC Health Serv Res 2013;13:119 [FREE Full text] [doi: 10.1186/1472-6963-13-119] [Medline: 23537350]
53. Zhang Z, Sarcevic A, Joy K, Ozkaynak M, Adelgais K. User needs and challenges in information sharing between pre-hospital and hospital emergency care providers. In: Proceedings of the American Medical Informatics Association Annual Symposium. 2021 Presented at: AMIA ’21; October 30-November 3, 2021; San Diego, CA.
54. Metatla O, Bryan-Kinns N, Stockman T. "I Hear You": understanding awareness information exchange in an audio-only workspace. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, Canada p. 1-13. [doi: 10.1145/3173574.3174120]
55. Follmann A, Ohligs M, Hochhausen N, Beckers SK, Rossaint R, Czaplik M. Technical support by smart glasses during a mass casualty incident: a randomized controlled simulation trial on technically assisted triage and telemedical app use in disaster medicine. J Med Internet Res 2019;21(1):e11939 [FREE Full text] [doi: 10.2196/11939] [Medline: 30609988]
56. First responder network authority. FirstNet Authority. URL: https://firstnet.gov/about [accessed 2021-05-25]
57. Gupta S, Boehme J, Manser K, Dewar J, Miller A, Siddiqui G, et al. Does wearable medical technology with video recording capability add value to on-call surgical evaluations? Surg Innov 2016;23(5):498-504. [doi: 10.1177/1553350616656278] [Medline: 27335083]
58. Zhang Z, Joy K, Upadhyayula P, Ozkaynak M, Harris R, Adelgais K. Data work and decision making in emergency medical services: a distributed cognition perspective. Proc ACM Hum-Comput Interact 2021;5(CSCW2):1-32. [doi: 10.1145/3479500]
59. Odenheimer S, Goyal D, Jones VG, Rosenblum R, Ho L, Chan AS. Patient acceptance of remote scribing powered by Google glass in outpatient dermatology: cross-sectional study. J Med Internet Res 2018;20(6):e10762 [FREE Full text] [doi: 10.2196/10762] [Medline: 29929947]

60. Aldaz G, Shluzas LA, Pickham D, Eris O, Sadler J, Joshi S, et al. Hands-free image capture, data tagging and transfer using Google glass: a pilot study for improved wound care management. PLoS One 2015;10(4):e0121179 [FREE Full text] [doi: 10.1371/journal.pone.0121179] [Medline: 25902061]
61. Maruri HA, Lopez-Meyer P, Huang J, Beltman WM, Nachman L, Lu H. V-Speech: noise-robust speech capturing glasses using vibration sensors. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018;2(4):1-23. [doi: 10.1145/3287058]
62. Boillat T, Grantcharov P, Rivas H. Increasing completion rate and benefits of checklists: prospective evaluation of surgical safety checklists with smart glasses. JMIR Mhealth Uhealth 2019;7(4):e13447 [FREE Full text] [doi: 10.2196/13447] [Medline: 31033451]
63. Mentler T, Berndt H, Herczeg M. Optical head-mounted displays for medical professionals: cognition-supporting human-computer interaction design. In: Proceedings of the European Conference on Cognitive Ergonomics. 2016 Presented at: ECCE '16; September 5-8, 2016; Nottingham, UK p. 1-8. [doi: 10.1145/2970930.2970957]
64. Bosk CL, Dixon-Woods M, Goeschel CA, Pronovost PJ. Reality check for checklists. Lancet 2009;374(9688):444-445. [doi: 10.1016/s0140-6736(09)61440-9] [Medline: 19681190]
65. Sparks EA, Wehbe-Janek H, Johnson RL, Smythe WR, Papaconstantinou HT. Surgical safety checklist compliance: a job done poorly!. J Am Coll Surg 2013;217(5):867-73.e1-3. [doi: 10.1016/j.jamcollsurg.2013.07.393] [Medline: 23973104]
66. Koelle M, Ananthanarayan S, Czupalla S, Heuten W, Boll S. Your smart glasses' camera bothers me!: exploring opt-in and opt-out gestures for privacy mediation. In: Proceedings of the 10th Nordic Conference on Human-Computer Interaction. 2018 Presented at: NordiCHI '18; September 29-October 3, 2018; Oslo, Norway p. 473-481. [doi: 10.1145/3240167.3240174]
67. Cobus V, Busse S, Heuten W. Google glass++: evaluating multimodal alarms on Google glass. In: Proceedings of Mensch und Computer 2019. 2019 Presented at: MuC'19; September 8-11, 2019; Hamburg, Germany p. 795-799. [doi: 10.1145/3340764.3344910]
68. Lateef F. The practice of telemedicine: medicolegal and ethical issues. Ethics Med 2011;27(1):17.

Abbreviations

ED: emergency department
EHR: electronic health record
EMS: emergency medical services
HIPAA: Health Insurance Portability and Accountability Act
MCI: mass casualty incident

Edited by A Kushniruk; submitted 01.06.21; peer-reviewed by A Follmann, S Bhattacharjee, KC Wong; comments to author 02.08.21; revised version received 23.11.21; accepted 07.12.21; published 28.02.22.

Please cite as:
Zhang Z, Joy K, Harris R, Ozkaynak M, Adelgais K, Munjal K
Applications and User Perceptions of Smart Glasses in Emergency Medical Services: Semistructured Interview Study
JMIR Hum Factors 2022;9(1):e30883
URL: https://humanfactors.jmir.org/2022/1/e30883
doi:10.2196/30883
PMID:35225816

Original Paper

Supporting Management of Noncommunicable Diseases With Mobile Health (mHealth) Apps: Experimental Study

Neta Kela[1], DPhil; Eleanor Eytam[1], DPhil; Adi Katz[1], DPhil

Shamoon College of Engineering, Ashdod, Israel

Corresponding Author:
Neta Kela, DPhil
Shamoon College of Engineering
84 Jabotinsky St
Ashdod, 77245
Israel
Phone: 972 088519309
Email: neta.kela@gmail.com

Abstract

Background: Noncommunicable diseases (NCDs) are the leading global health problem in this century and are the principal causes of death and health care spending worldwide. Mobile health (mHealth) apps can help manage and prevent NCDs if people are willing to use them as supportive tools. Still, many people are reluctant to adopt these technologies. Implementing new apps could result in earlier intervention for many health conditions, preventing more serious complications.

Objective: This research project aimed to test the factors that facilitate the adoption of mHealth apps by users with NCDs. We focused on determining, first, what user interface (UI) qualities and complexity levels appeal to users in evaluating mHealth apps. We also wanted to determine whether people prefer that the data collected by an mHealth app be analyzed using a physician or an artificial intelligence (AI) algorithm. The contribution of this work is both theoretical and practical. We examined users' considerations when adopting mHealth apps that promote healthy lifestyles and helped them manage their NCDs. Our results can also help direct mHealth app UI designers to focus on the most appealing aspects of our findings.

Methods: A total of 347 respondents volunteered to rate 3 models of mHealth apps based on 16 items that measured instrumentality, aesthetics, and symbolism. Respondents rated each model after reading 1 of 2 different scenarios. In one scenario, a physician analyzed the data, whereas, in the other, the data were analyzed by an AI algorithm. These scenarios tested the degree of trust people placed in AI algorithms versus the "human touch" of a human physician regarding analyzing data collected by an mHealth app.

Results: As shown by the responses, the involvement of a human physician in the application had a significant effect ($P<.001$) on the perceived instrumentality of the simple model. The complex model with more controls was rated significantly more aesthetic when associated with a physician performing data analysis rather than an AI algorithm ($P=.03$).

Conclusions: Generally, when participants found a human touch in the mHealth app (connection to a human physician who they assumed would analyze their data), they judged the app more favorably. Simple models were evaluated more positively than complex ones, and aesthetics and symbolism were salient predictors of preference. These trends suggest that designers and developers of mHealth apps should keep the designs simple and pay special attention to aesthetics and symbolic value.

(JMIR Hum Factors 2022;9(1):e28697) doi:10.2196/28697

KEYWORDS

mHealth; digital health; instrumentality; aesthetics; symbolic value; preference

Introduction

Background

Chronic diseases, known as noncommunicable diseases (NCDs), are the leading global health problem of this century [1]. According to the World Health Organization, these include cardiovascular diseases, cancer, chronic respiratory diseases, and diabetes mellitus [2]. These diseases are the principal causes of death and health care spending worldwide and are significant causes of poverty, which hinders economic development [3]. Ampofo and Boateng [4] suggested that, by 2030, the prevalence of obesity and diabetes will reach a peak in many countries. In addition, 20 million Americans are expected to have a history

of cancer by 2026, an increase that coincides with the increasing prevalence of obesity [5-7].

Fortunately, many chronic diseases can be delayed until significantly later in life, or even totally prevented, if people adopt a healthy lifestyle [1]. The digital health revolution—advances in medical information technologies such as information storage, data analysis, and genetic information, together with sensors embedded in smartphones [8]—can help people maintain healthy routines and manage chronic ailments. Mobile health (mHealth) apps are software applications developed for use on small wireless computing devices such as smartphones and tablets [9,10]. These apps can potentially impact people's health conditions because most of the global population has access to a mobile cellular network [1], and most people who have that access frequently check their phones [11]. This very high engagement level with smartphones presents an opportunity for health-oriented mHealth apps to help people lead healthier lifestyles and manage NCDs. Due to the socially authoritative influence of such apps, results can be highly effective. Still, people are often reluctant to adopt these supportive technologies, especially when they are asymptomatic [12], even though delayed treatment and intervention may, in turn, cause the disease to become irreversible.

mHealth apps serve a wide range of functions, and when adopted, they can help users cope and manage NCDs. It is estimated that such apps could cut annual US health care costs by US $150 billion by 2026 [13]. Still, the focus of most research to date has been on the judgments of physicians (eg, [1,14,15]) in postadoptive evaluations (eg, [16]). Less attention has been given to the patient's perspective and willingness to adopt technology in pre-adoptive evaluations. Thus, this research project aimed to employ pre-use evaluations to explore factors that facilitate the adoption of mHealth apps by users who cope with NCDs.

The user interface (UI) is the first point of contact between a user and an application. Preference assigned to mHealth apps largely depends on the qualities of the UI (eg, [17,18]). Users are interested in how useful an app may be, its easy operation, and its aesthetics [18,19]. In addition, technology in health care often relies on artificial intelligence (AI). Large and complex data sets (ie, big data) are used to train algorithms to learn and improve their analysis to support decision-making [20,21]. There is great optimism that AI can substantially improve diagnostics, treatment, and support in managing NCDs [13]. Even so, while clinicians are often reluctant to trust AI [14], little research has related to the willingness of users to rely on it in managing their health. Thus, this study revolved around user perspectives in adopting mHealth apps. The study addressed 2 primary concerns. First, what UI qualities and level of complexity of mHealth apps appeal to users? Second, do users prefer their data to be analyzed by an actual physician or an AI algorithm? The contribution of this work lies within its focus on users' considerations when adopting mHealth apps that can help them manage their NCDs.

We organize this work as follows: First, we explore the theoretical background of using technology in health care. Then, we focus on the contribution of mHealth apps in managing NCDs, followed by addressing human-computer interaction issues. Next, we propose 2 dimensions that influence preference: the quality of the design and the method by which data are analyzed (human versus AI analysis). We then continue and describe the study's methodology and report the study results. Afterward, we discuss directions for future research, and finally, we state our conclusions.

Technology in the Service of Health

Building on the role computers have come to play as counselors and experts, technology can promote a healthy lifestyle [22]. People assume that these "authorities" are intelligent and expect them to make suggestions and provide helpful information. Apps for healthier living that emphasize behavioral change and self-management have a high potential to help users achieve necessary lifestyle changes [23,24]. Two prominent examples of the effect of technology on health care are the Health Machine [23], which implements persuasion techniques to counter obesity and diabetes, and the Personal Aerobics Trainer, a virtual fitness trainer [25]. Technology can also support physicians in their daily tasks. For example, telemedicine enables physicians to diagnose and treat patients from afar. This technique dramatically reduces health care costs while creating a comfortable and safe treatment milieu [26]. Increased computer processing speed, the availability of large data sets, and a pool of talented AI developers have enabled the rapid development of technology in health care [27]. Moreover, AI algorithms often perform better than humans in an assortment of tasks [13].

mHealth Apps Supporting NCDs

Because smartphones are ideally suited for collecting medical data through features such as their camera, microphone, touch-sensitive screen, and accelerometer, the use of mHealth apps is increasing exponentially throughout the world [28]. This simple and socially acceptable means of collecting behavioral and physiological information [29] can support various health conditions [30]. mHealth apps can help monitor health indicators (eg, heart rate, blood pressure). They can also support people (eg, patients with diabetes or dementia) by monitoring their illnesses (eg, chronic obstructive pulmonary disease, hypertension, diabetes, dementia) [29,31] and their caregivers (eg, physicians, nurses). They accomplish this by providing education, synchronizing records, monitoring medications, or providing access to patient information (see [32] for an extensive overview of various categories of mHealth apps). These apps can serve both general and specific purposes. Although some focus on particular health dimensions (eg, diet and physical activity), others enable personal management of well-being by monitoring a diverse range of daily behaviors with broad health-related consequences. Apps can also be managed by operating on a "manual-automatic" scale, whose extreme ends are manually activated by users on one side and fully automatic on the other [8]. Many mobile, location-based exercise apps harness the power of gamification principles on GPS-enabled smartphones [33,34]. In some of these, augmented reality turns the real world into a "game map" or playground where users play while keeping fit [35,36].

Role of the Physician in Digital Health

Even though recent surveys indicate that more than one-third of American doctors recommend that their patients use a health or medical app [37], health care has been slow to incorporate technological advancements in clinical practice [38]. The most prominent reason for this is that the physician's role is undergoing tremendous change [39]. Since the earliest days of professional health care, doctors represented authority and knowledge and maintained responsibility for patient treatment. This traditional role is now shifting as patients can look up their symptoms on the internet and review the opinions of others regarding the best approach for treatment. Thus, the doctor no longer is the sole medical decision maker but becomes a vital member of a health care team [39,40].

Still, although the physician's role is changing, trust remains an essential and fundamental aspect of medical treatment [41]. In many cases, trust in the physician often plays a substantial part in patient recovery. A caring and competent physician increases this trust [40]. Perhaps this is why people turn to their doctor even though AI can provide many benefits [42]. Physicians still play a crucial role in guiding patients and helping them understand the information they encounter [40]. Because there is little regulation of medical apps or information on the internet, patients need this guidance from the doctor. In these situations, mHealth apps could be more beneficial if, before direct access to a doctor, trained AI bots can qualify whether specific symptoms warrant an actual visit [13], provided, of course, that people are willing to use them.

Persuasive Technologies

Persuasive technologies are interactive systems designed to foster behavioral and attitudinal changes [35]. mHealth apps are technologies in which persuasive design could be beneficial, motivating people toward healthy behaviors [42-44]. According to the Fogg Behavioral Model, one of the 3 motivators for a behavior to change is social acceptance. Most people are highly motivated to do things to further this acceptance. Marcus [23] suggested that social interaction has an important impact on behavioral change. For example, people on Facebook are significantly driven by a desire for social acceptance, which is why they share pictures, beliefs, and experiences. Given that people manage their image on social media platforms, how others perceive them also seems relevant to their health habits.

Further, just as social networking sites offer platforms to share accomplishments and foster collaboration and social support [45], it is reasonable to expect that mHealth apps could likewise provide a platform for collaborating, sharing, and receiving support in the area of health activities. Indeed, one previous study showed that creating a mobile virtual community for overweight individuals allowed them to receive social support, advice, and emotional encouragement [46]. This importance of social presence and symbolism aligns with Maslow's well-known theory of human motivation and needs [47]. The fundamental needs for *belonging and love* can be satisfied by health apps through being able to share health-related experiences with friends and family members, receive their support, and socially communicate a healthy image. In addition, the need for *esteem* can be met by displaying health-related accomplishments (such as weight and step count).

Factors Affecting the Intention to Use mHealth Apps

Salgado et al [48] recently reviewed various studies about mobile technology solutions to address health care challenges. Following their review, they suggested that the presence of a chronic health condition predicates an impact on the acceptance of mHealth technology. Huong and Long [49] showed that the intention to use mHealth apps is affected mainly by mHealth literacy, perceived usefulness, and perceived ease of use. The concept of mHealth literacy is drawn from the compatibility suggested in Rogers' classical theory of innovation diffusion [50]. Compatibility refers to the level at which a product is compatible with a potential user's past experiences and beliefs. Both compatibility and literacy suggest that the more technologically literate users are, the more likely they will find innovation compatible with their values and beliefs.

Because not all users have the same level of technological literacy, the app itself should appeal to users. *Perceived usefulness* and *perceived ease of use* are constructs of the technology acceptance model, a well-known model for understanding the intention of utilizing innovative technology [51]. Perceived usefulness refers to the degree to which individuals believe that using a specific technology can improve their task performance. Perceived ease of use is the subjective belief that the product, when used, does not require significant physical or mental effort. The higher these two constructs are rated, the greater the intention to use the product [46,48]. Paying attention to these constructs in an app's design can encourage use of these technological tools.

Qualities of User Interfaces and Effect on Preference

Three major product qualities are essential in evaluating an interactive product: instrumentality, aesthetics, and symbolism [52]. Instrumentality relates to how a product fulfills the practical needs of promoting the users' goals through usability. Instrumentality is an aggregate of perceived usefulness and perceived ease of use [53]. Aesthetics revolve around the sensual effect the product has on the user, eliciting an emotional reaction of, for example, tranquility, confidence, pleasantness, or frustration. Symbolism refers to the associations that the product produces and the meanings it communicates, regardless of its pragmatic goals. The effect of each of these qualities on product preference is mediated by the role of the users [18], their personal characteristics [19], and the product itself [17]. Eytam et al [18] found that the visual simplicity or complexity of a UI, as reflected by the number of its controls, influences judgments of instrumentality and aesthetics. Still, aesthetics is a consistent predictor of preference of UIs for both simple and complex designs. Symbolism is found to influence decisions about a product's characteristics. It is a salient predictor of the perceived creativity of product UIs, regardless of their complexity level [17].

Research Hypotheses

Our model postulates that 2 significant elements influence preferences for mHealth apps. The first of these includes product qualities, namely, instrumentality, aesthetics, and symbolism

[52], as reflected by the number of controls in the UI (needed to operate it) [18]. The second focuses on whether a physician or an AI algorithm performs the data analysis of information collected by the app. The following hypotheses explain how these 2 dimensions may affect preferences for different models of an mHealth app.

Because users are more likely to trust a human physician [40,41], we expect an app with a physician intervention would be rated higher in instrumental value than the same application backed by the support of an AI algorithm. Similarly, because apps presenting an excessive number of controls are reported in the literature to be more complicated to use [53,54], we expected that the number of controls will affect instrumentality ratings. Having fewer controls was expected to increase instrumental value regardless of data analysis mode [18]. Thus, H1a was that "Instrumentality judgments of mHealth apps should be higher when data are analyzed by a human physician versus an AI algorithm." H1b was that "Instrumentality judgments of mHealth apps should be higher when there are few versus many controls, regardless of mHealth data analysis mode."

Because the data analysis process is embedded in the system and is not reflected in the design, the presence of a human physician or an AI algorithm to analyze the data was not expected to affect noninstrumental judgments [17] that revolve around user delight and satisfaction [55]. Therefore, H2a was that "Aesthetic judgments of mHealth apps should be similar when data are analyzed either by a human physician or an AI algorithm." H2b was that "Symbolism judgments of mHealth apps should be similar when data are analyzed either by a human physician or an AI algorithm."

Data analysis is a pragmatic characteristic of the application that is not reflected in the design [17]. Therefore, we expected instrumentality to be a salient predictor of product preference for applications backed by a human physician and believed to have greater instrumental value. Because noninstrumental attributes are reported in the literature as salient predictors of preference [18,19], we expected aesthetics and symbolism in both application types (with a human physician and with an AI algorithm) to be salient predictors of mHealth apps. Therefore, H3a was that "Instrumentality should be a salient predictor of preference variance for apps that engage a human physician versus an AI algorithm." H3b was that "Aesthetics should be a salient predictor of preference variance for apps that engage a human physician and those that engage an AI algorithm." H3c was that "Symbolism should be a salient predictor of preference variance for apps that engage a human physician and those that engage an AI algorithm."

Because traditional health care is characterized by personal contact (human touch) between a patient and caregiver [14,40], we expected that preferences for apps that engage a human physician would be higher than those that rely on automatic AI analysis. H4 was that "Preference is higher for apps that engage a human physician versus an AI algorithm."

Methods

User Evaluation

In this research project, we conducted a user evaluation of 3 key UI features of mHealth apps: instrumentality, aesthetics, and symbolism. We compared user responses to descriptions of apps that use AI to analyze the data collected, while the app also had a physician available to analyze the same data remotely. To test if there was a difference in user preferences, we asked the 2 respondent groups to respond in writing to a questionnaire used to rate 3 different models of an mHealth app. These models differed in the number of their controls. Although each group of respondents evaluated the same models, before each group began to complete the same questionnaire, the members of one group received a different scenario than the members of the other group. The first group was told that a physician would examine the data received from the mHealth app (hereafter referred to as the physician, or doctor, scenario). The second group was told that data received by the mHealth app would be analyzed by a very accurate AI algorithm (hereafter referred to as the robot, or bot, scenario). Thus, the research was designed as a between-dimensions (2 scenarios/app descriptions) and a within-dimensions (3 models/stimuli) experiment.

Sample

There were 347 respondents who took part in the study (mean age 29.12, SD 9.20, range 15-86 years; gender: 198/347, 57.1% female). Respondents were volunteers recruited by students taking a data analysis course at an engineering college.

Stimuli

The mHealth app features were designed by students participating in a UI course. The features the students were asked to create had to fit 1 of 3 themes: frequently used mHealth features, health indicators, and social-oriented features. The final designs were refined by 3 judges (2 human-computer interaction specialists and 1 biologist). The final stimuli involved 3 models: The first model was simple—with a 4-control design including frequently used mHealth features. The second model was medium—with an 8-control design presenting added health indicators. The third model was complex—with a 12-control design that included added social-oriented features (Multimedia Appendix 1 presents the 3 models of the application). Each control represented a different feature commonly used in well-being (eg, an iPhone health app) and diet-supporting applications.

Measures

We borrowed 16 items measuring instrumentality, aesthetics, and symbolism (Multimedia Appendix 2) from the human-computer interaction literature [17].

Manipulation

In order to manipulate the use scenario (doctor versus robot), a short introduction preceded the questionnaire and introduced either the human doctor or an automatic AI algorithm (See Multimedia Appendix 3 for each introduction).

Procedure

Respondents were randomly assigned to 1 of 2 groups: a group presented with a doctor scenario (n=159) and a group presented with a robot scenario (n=188). The members of each group read the scenario preface for their group only, before anonymously completing the questionnaire. The illustration of each model was presented 4 times, each time with a different set of 4 randomly chosen items, to control for possible consistency effects. To conclude the study, respondents were asked to rate their preference regarding each design on a Likert scale (1-7) and choose their favorite application design.

Ethics Approval

The Shamoon College of Engineering IRB (ethics committee) approved the research project (review 12), including the experimental task, the testing procedure, and the collection of data.

Results

An analysis of standard residuals was carried out on the data to identify any outliers. The analysis results indicated that 22 (6.3%) of the total sample (347 respondents) needed to be removed because they responded similarly to all different items for the 3 designs tested. Of the respondents, before rating the different designs, 144 read the doctor-scenario description, and 181 read the robot-scenario description. Responses to items describing attributes of design illustrations were subjected to exploratory factor analysis. Following Rafaeli and Vilnai-Yavetz [52], our theoretical model assumed 3 distinct factors corresponding to the product qualities of instrumentality, aesthetics, and symbolism. Accordingly, 3 factors were specified for retention. Maximum likelihood estimation and oblique rotation (direct oblimin with Kaiser normalization) were applied separately to the data for each model tested (ie, Models 1, 2, and 3). Items were loaded on 3 distinct factors for all models (Multimedia Appendix 4 presents factor loadings of items for the 3 models tested). The 3 factors explained 76% to 77% of the variance in each of the 3 analyses. The items of each attribute were averaged to create scale scores. Cronbach alpha reliabilities were calculated for the attributes of each illustration in each group. All scales had adequate reliabilities (between 0.88 and 0.93) in all conditions.

In general, the correlations between the scales for the 3 models tested were between 0.62 and 0.73, which is in line with previous studies [56-58]. For Models 1 and 2, the correlations were not excessive in any of the conditions, an outcome that indicates reasonable discriminability for all 3 attribute ratings that occurred. For Model 3, the correlations between attribute ratings exceeded 0.70. These correlations may indicate that it was too difficult to differentiate between the different qualities with too many controls. Table 1 presents the correlations and reliabilities for each scale in each product condition.

Table 1. Correlations and reliabilities for each scale in each condition (n=325).

Level	Model 1			Model 2			Model 3		
	Instrumentality	Aesthetics	Symbolism	Instrumentality	Aesthetics	Symbolism	Instrumentality	Aesthetics	Symbolism
Instrumentality	0.91[a]	0.62	0.67	0.90[a]	0.66	0.68	0.91[a]	0.73	0.71
Aesthetics	0.62	0.92[a]	0.69	0.66	0.93[a]	0.69	0.73	0.92[a]	0.70
Symbolism	0.67	0.69	0.88[a]	0.68	0.69	0.88[a]	0.71	0.70	0.90[a]

[a]Reliability.

A series of mixed-design analysis of variance studies were conducted with *product* (doctor versus robot) as a between-groups factor and *model* (1, 2, or 3) as a within-subjects factor. *Instrumentality*, *aesthetics*, *symbolism*, and *preference* were the dependent variables. The Mauchly test indicated that the assumption of sphericity had been violated (instrumentality: χ^2_2=130.28, *P*<.001; aesthetics: χ^2_2=51.74, *P*<.001; symbolism: χ^2_2=48.54, *P*<.001; preference: χ^2_2=75.29, *P*<.001). Therefore, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (instrumentality: ε=0.76; aesthetics: ε=0.88; symbolism: ε=0.88; preference: ε=0.83). All pairwise comparisons used the Bonferroni correction for multiple tests. Figures 1-4 detail ratings for product attributes for the 2 product conditions tested.

Figure 1. Average instrumentality ratings of doctor versus robot based on model. AI: artificial intelligence.

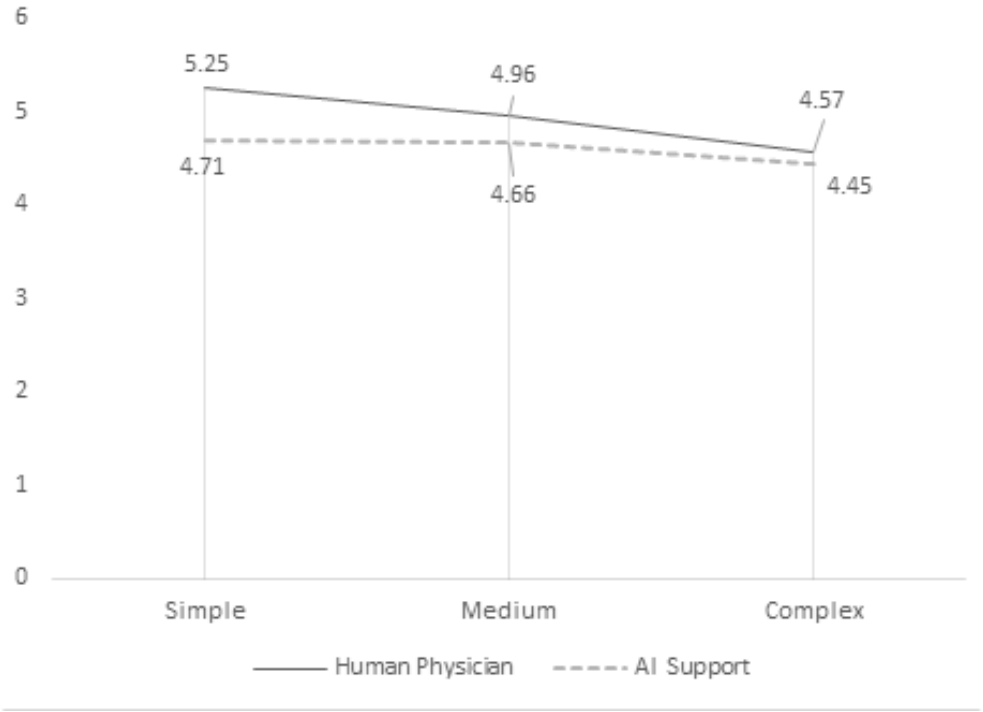


Figure 2. Average aesthetics ratings of doctor versus robot based on model. AI: artificial intelligence.

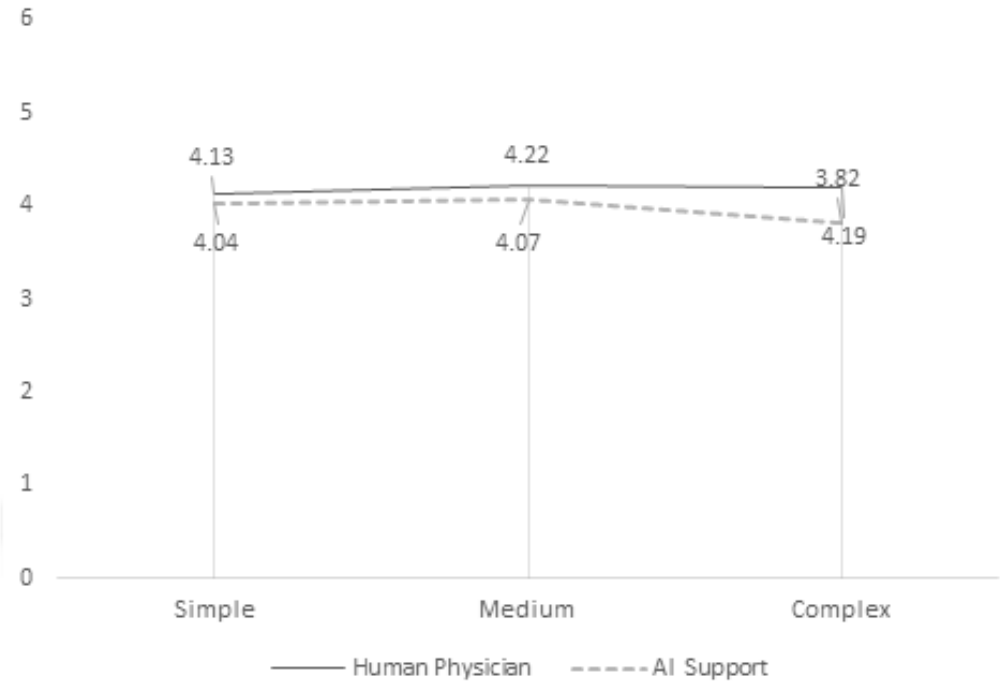


Figure 3. Average symbolism ratings of doctor versus robot based on model. AI: artificial intelligence.

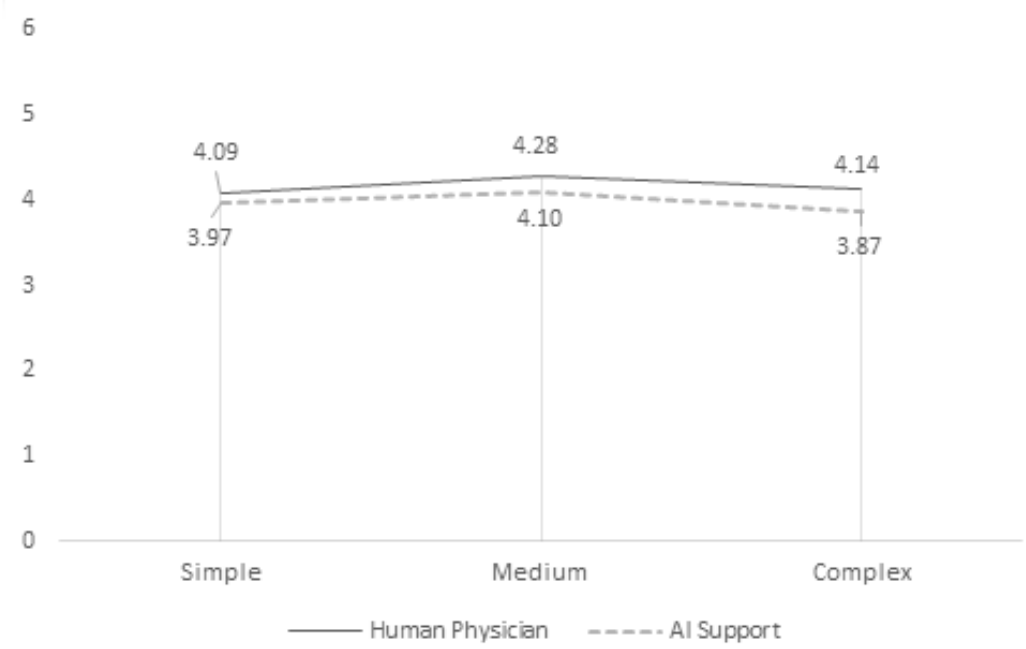
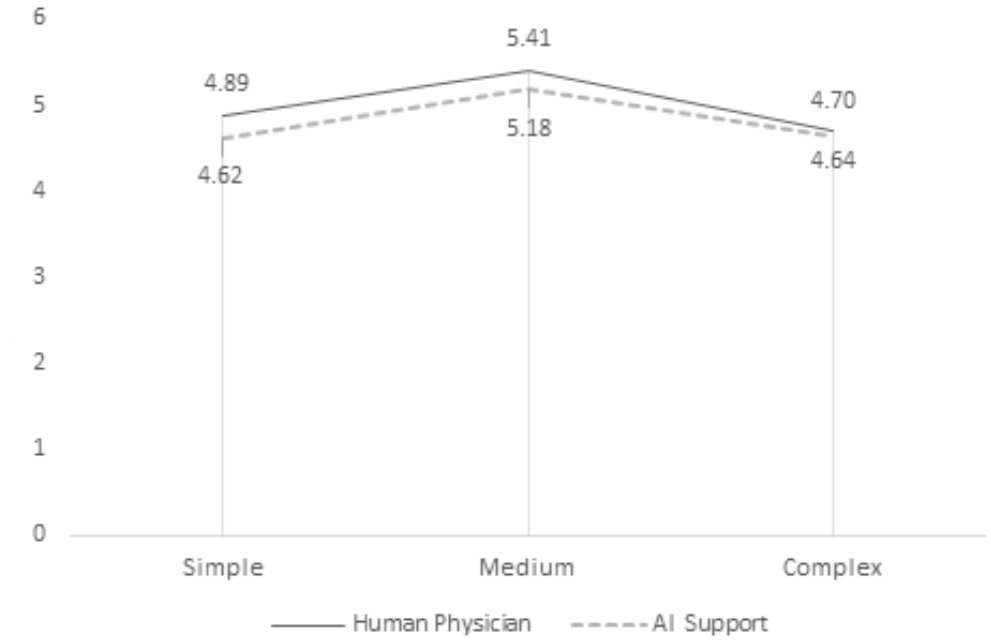


Figure 4. Average preference ratings of doctor versus robot based on model. AI: artificial intelligence.



Model had a significant effect on instrumentality, aesthetics, symbolism, and preference ratings. There was a positive relationship between model and instrumentality ratings. The Model 1 (doctor: mean 5.25, SD 1.57; robot: mean 4.71, SD 1.59) and Model 2 (doctor: mean 4.96, SD 1.56; robot: mean 4.66, SD 1.50) designs were rated as significantly more instrumental than that of Model 3 (doctor: mean 4.57, SD 1.72; robot: mean 4.45, SD 1.71) in both conditions (both comparisons, $P<.001$). Similar trends were also reported in studies that examined the effect of model choice on judgments of instrumentality [18,19]. There was a significant difference between aesthetic ratings of Model 2 (mean 4.07, SD 1.60) versus Model 3 (mean 3.82, SD 1.60) but only in the robot condition ($P=.004$). There was a significant difference between the symbolism ratings of Model 2 (doctor: mean 4.28, SD 1.60; robot: mean 4.10, SD 1.48) compared with those of Model 3 (doctor: mean 4.14, SD 1.65; robot: mean 3.87, SD 1.55; $P=.001$). Finally, there was a significant difference between preference ratings for Model 1 (doctor: mean 4.89, SD 1.77; robot: mean 4.62, SD 1.88) and Model 3 (doctor: mean 4.70, SD 1.8760; robot: mean 4.64, SD 1.83) compared with those of Model 2 (doctor: mean 5.41, SD 1.32; robot: mean 5.18, SD 1.36; $P<.001$).

Group had a significant effect on the instrumentality ratings of Model 1 (doctor: mean 5.25, SD 1.57; robot: mean 4.71, SD 1.59; $P<.001$) and on the aesthetics ratings of Model 3 (doctor: mean 4.19, SD 1.70; robot: mean 3.82, SD 1.60; $P=.03$). The interaction effect was significant for the *instrumentality*

($F_{2,646}$=8.05, *P*=.001) and *aesthetics* rating ($F_{2,646}$=4.19, *P*=.02), indicating that the effect of the virtual presence of a human physician was greater in judgments of instrumentality of Model 1 and in judgments of aesthetics of Model 3.

We conducted separate regression analyses for each model (1, 2, and 3), with preference as the dependent variable and product attributes (instrumentality, aesthetics, and symbolism) as the predictors. Tests to see if the data met the assumption of collinearity indicated that multicollinearity was not a concern for any of the analyses conducted (tolerance >.2) [59]. The results of the regression are presented in Table 2 and Table 3. In the doctor condition, independent variables (product attributes) accounted for 13% to 18% of the preference variance in the Models 1, 2, and 3 analyses. When reading the doctor scenario before evaluating Model 1, respondents considered all product attributes as salient. When reading the robot scenario before evaluating Model 1, respondents considered only aesthetics as a salient predictor for preference. When preparing to evaluate Model 2, respondents who read either scenario found one noninstrumental attribute as salient (although they found different attributes—symbolism for the doctor scenario and aesthetics for the robot). The 2 different scenarios did not influence the importance of any of the 3 product attributes in any significant manner. Probably adding many controls to the model, as in Model 3, brings about different considerations that we did not measure in this research.

To test the percentage of respondents preferring different designs after reading each scenario, we used a Z-ratio test (based on the calculator in [60]). The Z-ratios for proportions of design choice frequencies and group were not significant for the simple design (Z=0.455, *P*=.65), medium design (Z=–0.210,*P*=.48), and Model 3 (Z=–0.211, *P*=.83). The frequencies of the choice of design are depicted in Table 4.

Table 2. Preference model standardized regression coefficients (doctor [n=144] scenario).

Condition	Doctor model 1 (R^2=0.18)				Doctor model 2 (R^2=0.13)				Doctor model 3 (R^2=0.17)			
	Beta	SE	*P* value	*F* (df)	Beta	SE	*P* value	*F* (df)	Beta	SE	*P* value	*F* (df)
Instrumentality	–0.34	0.13	*P*=.004	10.40 (3,140)	0.13	0.11	*P*=.34	6.87 (3,140)	0.20	0.16	*P*=.17	9.68 (3,140)
Aesthetics	0.32	0.11	*P*=.004		0.02	0.10	*P*=.88		0.14	0.14	*P*=.29	
Symbolism	0.38	0.11	*P*=.001		0.25	.09	*P*=.02		0.12	0.13	*P*=.28	

Table 3. Preference model standardized regression coefficients (robot [n=181] scenario).

Condition	Robot model 1 (R^2=0.27)				Robot model 2 (R^2=0.16)				Robot model 3 (R^2=0.11)			
	Beta	SE	*P* value	*F* (df)	Beta	SE	*P* value	*F* (df)	Beta	SE	*P* value	*F* (df)
Instrumentality	–0.10	0.11	*P*=.25	21.63 (3,177)	–0.06	0.09	*P*=.55	11.38 (3,177)	0.04	0.11	*P*=.71	7.37 (3,177)
Aesthetics	0.48	0.12	*P*=.001		0.01	0.09	*P*=.95		0.13	0.13	*P*=.26	
Symbolism	0.12	0.14	*P*=.33		0.43	0.12	*P*=.001		0.20	0.14	*P*=.10	

Table 4. Choices of design (doctor versus robot scenario; total n=325).

Scenario	Model 1 (n=84), n	Model 2 (n=108), n	Model 3 (n=133), n
Doctor (n=144)	39	47	58
Robot (n=181)	45	61	75

Discussion

Principal Findings

This research project investigated user preferences for mHealth apps. We sought to facilitate the acceptability of such technology in health care provision, which would lead to more frequent and productive use of these apps. In general, when a human touch was present in the analysis, that is, when the respondents thought a physician would analyze the data collected by the mHealth app, ratings of both instrumentality and aesthetics were higher than the scenario in which they thought AI would analyze their data. These overall higher ratings can be explained by trust. Previous studies reported that people do not trust AI-based technology in health care as much as they do their doctors (eg, [15,42]). A human physician increases the sense of connectedness to a knowledgeable, caring health care professional [41].

In contrast, an AI algorithm works as a “black box”—a metaphor suggesting that, because people do not know how they produce their outputs, they have less trust in them [15]. Vo et al [61] reviewed 43 qualitative studies of patients’ perceptions of mHealth. They found that patients appreciated communicating directly with health care professionals and providers because they could receive responses to their concerns from a person who cared. Patients with chronic ailments reported that they want to share their health records with their physicians between clinic visits [62] because of their need for a relationship with the caregiver [41]. Table 5 summarizes the research hypotheses.

Table 5. Research hypotheses.

Hypothesis number	Hypothesis description	Model 1[a]	Model 2[b]	Model 3[c]	Table or figure
H1[a]	Instrumentality judgments of mHealth[d] apps should be higher when data are analyzed by a human physician versus an AI[e] algorithm.	√	X	X	Figure 1
H1[b]	Instrumentality judgments of mHealth apps should be higher when there are few versus many controls, regardless of mHealth data analysis mode.	√	√	√	Figure 1
H2[a]	Aesthetic judgments of mHealth apps should be similar when data are analyzed either by a human physician or an AI algorithm.	√	√	X	Figure 2
H2[b]	Symbolism judgments of mHealth apps should be similar when data are analyzed either by a human physician or an AI algorithm.	√	√	√	Figure 3
H3[a]	Instrumentality should be a salient predictor of preference variance for apps that engage a human physician versus an AI algorithm.	√	X	X	Table 2
H3[b]	Aesthetics should be a salient predictor of preference variance for apps that engage a human physician and those that engage an AI algorithm.	√	X	X	Table 2
H3[c]	Symbolism should be a salient predictor of preference variance for apps that engage a human physician and those that engage an AI algorithm.	X	√	X	Table 2
H4	Preference is higher for apps that engage a human physician versus an AI algorithm.	X	X	X	Figure 4

[a]4 controls in the design.

[b]8 controls in the design.

[c]12 controls in the design.

[d]mHealth: mobile health.

[e]AI: artificial intelligence.

The simplest model (Model 1 with 4 controls) was judged the most instrumental among the 3 models tested. Predictably, the most complex model (Model 3 with 12 controls) was regarded as the least instrumental. This pattern of rating simplicity as providing high instrumentality has been noted in previous research [18,19]. Usability experts often advocate simplicity to promote a product's usability. They suggest that simple designs help people achieve their goals more efficiently and effectively because of their clarity and filtering out unnecessary features [63]. Hilliard et al [64] reported that chronically ill patients preferred apps that required minimal effort to input medical data or to set up scheduled alarms. In addition, respondents in our study, regardless of the scenario they read before responding in writing to the survey, preferred mid-level complexity (Model 2 with 8 controls). This preference for Model 2 hints at the idea that, while users do not want restricted functionality, they also do not want feature-laden apps [17,18].

The complex design was rated significantly more aesthetic when a human physician analyzed the data than an AI algorithm. Simplicity is often associated with beauty [63] and sophistication [65,66]. The effect of a human physician's involvement on aesthetic perceptions could be derived from a halo effect that made the overall impression of the application more positive in general because of this feature. Even so, previous research found that aesthetic websites enhance customer trust [67-69]. Perhaps this effect is also reversed, and confidence in a human physician's involvement in the app made it appear more aesthetic.

Noninstrumental qualities, namely, aesthetics and symbolism, were significant predictors of preference variance in both types of eHealth applications tested, which hints at the salient role of hedonic qualities in the evaluation of the app. Although potential users of mHealth apps have primarily utilitarian needs [55], users of technology products tend to stress hedonic motivations [56,57]. Eytam et al [18] noted that aesthetics is a consistent predictor of preference variance. The negative effect of instrumentality on product preference when a human physician is involved in data analysis may suggest that users' needs are not settled when their usability expectations are met but rather that they seek the hedonic benefits of the app.

This study explored how mHealth app qualities can affect the willingness of patients with NCDs to adopt these tools in their daily routine. Although it included the primary app qualities of instrumentality, aesthetics, and symbolism, it did not delve into the specific functions that patients look for in mHealth apps. That said, the literature suggests that specific functions such as connectedness to a support group through social media can promote mHealth apps [23]. Future studies should relate to particular features in these apps that can encourage willingness to adopt them. Specifically, future research should examine how widening the human touch in applications via connectedness to support groups may affect the acceptability of mHealth apps.

Conclusion

Our research model proposed 2 dimensions that influence app preference: design quality and the method of data analysis (ie,

by a human physician or AI algorithm analysis). We tested 3 application models to study these factors, each with a different number of controls for the various functions. Initially, we hypothesized that human touch in the application in the form of an assumed analysis of the data by a human physician would be perceived as more attractive than one automatically analyzed by an AI algorithm. The involvement of a physician in the application had a significant effect on the perceived instrumentality only for the simple design; however, physician involvement did not affect preference for an app. This lack of ability to affect preference is probably because judgments of the noninstrumental qualities—aesthetics and symbolism—which are the significant predictors of preference variance, were unaffected by how the data were analyzed. Overall, our findings show that mHealth adoption can be facilitated when the complexity of the design is restricted, when hedonic qualities of the design are attended to, and when human touch with a physician is taken into account. Because previous research suggests that aesthetics enhance trust in technology, investing in the aesthetics of mHealth apps would be a wise strategy to promote adoption by potential users.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Stimuli.
[DOCX File , 451 KB - humanfactors_v9i1e28697_app1.docx]

Multimedia Appendix 2
Items.
[DOCX File , 16 KB - humanfactors_v9i1e28697_app2.docx]

Multimedia Appendix 3
Scenarios.
[DOCX File , 16 KB - humanfactors_v9i1e28697_app3.docx]

Multimedia Appendix 4
Factor loadings.
[DOCX File , 15 KB - humanfactors_v9i1e28697_app4.docx]

References

1. Zykin LF. [The absence of differences between El Tor cholera and classic cholera]. Zh Mikrobiol Epidemiol Immunobiol 1992 Apr(4):76. [Medline: 1496887]
2. Sagner M, McNeil A, Puska P, Auffray C, Price ND, Hood L, et al. The P4 health spectrum - a predictive, preventive, personalized and participatory continuum for promoting healthspan. Prog Cardiovasc Dis 2017 Mar;59(5):506-521. [doi: 10.1016/j.pcad.2016.08.002] [Medline: 27546358]
3. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, American Heart Association Statistics Committee, Stroke Statistics Subcommittee. Executive summary: heart disease and stroke statistics--2016 update: a report from the American Heart Association. Circulation 2016 Jan 26;133(4):447-454. [doi: 10.1161/CIR.0000000000000366] [Medline: 26811276]
4. Ampofo AG, Boateng EB. Beyond 2020: Modelling obesity and diabetes prevalence. Diabetes Res Clin Pract 2020 Sep;167:108362. [doi: 10.1016/j.diabres.2020.108362] [Medline: 32758618]
5. Hong Y, Huo J, Desai R, Cardel M, Deshmukh AA. Excess costs and economic burden of obesity-related cancers in the United States. Value Health 2019 Dec;22(12):1378-1386 [FREE Full text] [doi: 10.1016/j.jval.2019.07.004] [Medline: 31806194]
6. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of Obesity Among Adults and Youth: United States, 2015–2016. Centers for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/nchs/products/databriefs/db288.htm [accessed 2022-01-27]
7. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: payer-and service-specific estimates. Health Aff (Millwood) 2009 Jan;28(5):w822-w831. [doi: 10.1377/hlthaff.28.5.w822] [Medline: 19635784]
8. Lane ND, Mohammod M, Lin M, Yang X, Lu H, Ali S, et al. Bewell: A smartphone application to monitor, model and promote well-being. 2011 Presented at: 5th International ICST Conference on Pervasive Computing Technologies for Healthcare; May 23-26, 2011; Dublin, Ireland. [doi: 10.4108/icst.pervasivehealth.2011.246161]
9. Weichbroth P. Usability of mobile applications: a systematic literature study. IEEE Access 2020;8:55563-55577. [doi: 10.1109/access.2020.2981892]

10. Wigmore I. Mobile app. Whatis.com. URL: https://whatis.techtarget.com/definition/mobile-app [accessed 2022-01-27]
11. O'Donnell D. Americans now check their phones up to 96 times per day: new study. Notebook Check. 2020 Mar 13. URL: https://www.notebookcheck.net/Americans-now-check-their-phones-up-to-96-times-per-day-new-study.457468.0.html [accessed 2022-01-27]
12. Lee M, Kang D, Yoon J, Shim S, Kim I, Oh D, et al. The difference in knowledge and attitudes of using mobile health applications between actual user and non-user among adults aged 50 and older. PLoS One 2020 Oct 27;15(10):e0241350 [FREE Full text] [doi: 10.1371/journal.pone.0241350] [Medline: 33108792]
13. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artificial Intelligence in Healthcare 2020:25-60 [FREE Full text] [doi: 10.1016/B978-0-12-818438-7.00002-2]
14. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. J Med Internet Res 2020 Jun 19;22(6):e15154 [FREE Full text] [doi: 10.2196/15154] [Medline: 32558657]
15. Gille F, Jobin A, Ienca M. What we talk about when we talk about trust: Theory of trust for AI in healthcare. Intelligence-Based Medicine 2020 Nov;1-2:100001. [doi: 10.1016/j.ibmed.2020.100001]
16. O'Connor Y, Andreev P, O'Reilly P. MHealth and perceived quality of care delivery: a conceptual model and validation. BMC Med Inform Decis Mak 2020 Feb 27;20(1):41 [FREE Full text] [doi: 10.1186/s12911-020-1049-8] [Medline: 32103746]
17. Eytam E. Effect of visual design on the evaluation of technology- vs. design-based novel interactive products. Interacting with Computers 2020;32(3):296-315 [FREE Full text] [doi: 10.1093/iwc/iwaa021]
18. Eytam E, Tractinsky N, Lowengart O. The paradox of simplicity: Effects of role on the preference and choice of product visual simplicity level. International Journal of Human-Computer Studies 2017 Sep;105:43-55. [doi: 10.1016/j.ijhcs.2017.04.001]
19. Eytam E, Lowengart O, Tractinsky N. Effects of visual simplicity in product design and individual differences in preference of interactive products. Rev Manag Sci 2020 Apr 23;15(5):1347-1389. [doi: 10.1007/s11846-020-00391-3]
20. Panesar A. Machine Learning and AI for Healthcare: Big Data for Improved Health Outcomes. Berkeley, CA: Apress; 2019.
21. Shahbaz M, Gao C, Zhai L, Shahzad F, Hu Y. Investigating the adoption of big data analytics in healthcare: the moderating role of resistance to change. J Big Data 2019 Jan 31;6(1):1. [doi: 10.1186/s40537-019-0170-y]
22. Fogg B. Computers as persuasive social actors. Persuasive Technology. 2003. URL: https://www.researchgate.net/publication/256867131_Computers_as_Persuasive_Social_Actors [accessed 2022-01-27]
23. Marcus A. The health machine. IDJ 2011 Sep 2;19(1):69-89. [doi: 10.1075/idj.19.1.09mar]
24. Edwards EA, Lumsden J, Rivas C, Steed L, Edwards LA, Thiyagarajan A, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. BMJ Open 2016 Oct 04;6(10):e012447 [FREE Full text] [doi: 10.1136/bmjopen-2016-012447] [Medline: 27707829]
25. Davis J, Bobick AF. Virtual PAT: A Virtual Personal Aerobics Trainer. Ohio State. 1998. URL: https://web.cse.ohio-state.edu/~davis.1719/Publications/pui98.pdf [accessed 2022-01-27]
26. Bashshur R, Doarn CR, Frenk JM, Kvedar JC, Woolliscroft JO. Telemedicine and the COVID-19 pandemic, lessons for the future. Telemed J E Health 2020 May 01;26(5):571-573. [doi: 10.1089/tmj.2020.29040.rb] [Medline: 32275485]
27. Lee KF. Ai Superpowers: China, Silicon Valley, and the New World Order. New York, NY: Harper Business; 2018.
28. Latif S, Rana R, Qadir J, Ali A, Imran MA, Younis MS. Mobile health in the developing world: review of literature and lessons from a case study. IEEE Access 2017;5:11540-11556. [doi: 10.1109/access.2017.2710800]
29. Saleh S, Farah A, Dimassi H, El Arnaout N, Constantin J, Osman M, et al. Using mobile health to enhance outcomes of noncommunicable diseases care in rural settings and refugee camps: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jul 13;6(7):e137 [FREE Full text] [doi: 10.2196/mhealth.8146] [Medline: 30006326]
30. Borycki E, Monkman H, Griffith J, Kushniruk A. Mobile usability testing in healthcare: methodological approaches. Stud Health Technol Inform 2015;216:338-342. [Medline: 26262067]
31. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. JMIR Mhealth Uhealth 2019 Jul 04;7(7):e11926 [FREE Full text] [doi: 10.2196/11926] [Medline: 31274112]
32. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014 Feb 05;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]
33. Tabi K, Randhawa AS, Choi F, Mithani Z, Albers F, Schnieder M, et al. Mobile apps for medication management: review and analysis. JMIR Mhealth Uhealth 2019 Sep 11;7(9):e13608 [FREE Full text] [doi: 10.2196/13608] [Medline: 31512580]
34. Yin AL, Hachuel D, Pollak JP, Scherl EJ, Estrin D. Digital health apps in the clinical care of inflammatory bowel disease: scoping review. J Med Internet Res 2019 Aug 19;21(8):e14630 [FREE Full text] [doi: 10.2196/14630] [Medline: 31429410]
35. Boulos MN, Yang SP. Exergames for health and fitness: the roles of GPS and geosocial apps. Int J Health Geogr 2013;12(1):18. [doi: 10.1186/1476-072x-12-18]
36. Freeman B, Chau J, Mihrshahi S. Why the public health sector couldn't create Pokémon Go. Public Health Res Pract 2017 Jul 26;27(3):2731724 [FREE Full text] [doi: 10.17061/phrp2731724] [Medline: 28765857]

37. Loh E. Medicine and the rise of the robots: a qualitative review of recent advances of artificial intelligence in health. leader 2018 Jun 01;2(2):59-63. [doi: 10.1136/leader-2018-000071]
38. Meskó B, Győrffy Z. The rise of the empowered physician in the digital health era: viewpoint. J Med Internet Res 2019 Mar 26;21(3):e12490 [FREE Full text] [doi: 10.2196/12490] [Medline: 30912758]
39. Meskó B, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. Mhealth 2017 Sep;3:38-38 [FREE Full text] [doi: 10.21037/mhealth.2017.08.07] [Medline: 29184890]
40. Thom DH, Hall MA, Pawlson LG. Measuring patients' trust in physicians when assessing quality of care. Health Aff (Millwood) 2004 Jul;23(4):124-132. [doi: 10.1377/hlthaff.23.4.124] [Medline: 15318572]
41. Yokoi R, Eguchi Y, Fujita T, Nakayachi K. Artificial intelligence is trusted less than a doctor in medical treatment decisions: influence of perceived care and value similarity. International Journal of Human–Computer Interaction 2020 Dec 26;37(10):981-990. [doi: 10.1080/10447318.2020.1861763]
42. Cheng R. Persuasion strategies for computers as persuasive technologies. Flylib. 2010. URL: https://flylib.com/books/en/2.438.1/computers_as_persuasive_tools.html [accessed 2022-01-27]
43. Intille SS. Ubiquitous computing technology for just-in-time motivation of behavior change. Stud Health Technol Inform 2004;107(Pt 2):1434-1437. [Medline: 15361052]
44. Kraft P, Drozd F, Olsen E. Digital Therapy: Addressing Willpower as Part of the Cognitive-Affective Processing System in the Service of Habit Change. In: Oinas-Kukkonen H, Hasle P, Harjumaa M, Segerståhl K, Øhrstrøm P, editors. Persuasive Technology. PERSUASIVE 2008. Lecture Notes in Computer Science, vol 5033. Berlin, Heidelberg: Springer; 2008:177-188.
45. Chan A, Kow R, Cheng JK. Adolescents' perceptions on smartphone applications (apps) for health management. JournalMTM 2017 Aug;6(2):47-55. [doi: 10.7309/jmtm.6.2.6]
46. Fukuoka Y, Kamitani E, Bonnet K, Lindgren T. Real-time social support through a mobile virtual community to improve healthy behavior in overweight and sedentary adults: a focus group analysis. J Med Internet Res 2011 Jul 14;13(3):e49 [FREE Full text] [doi: 10.2196/jmir.1770] [Medline: 21752785]
47. Maslow AH. A theory of human motivation. Psychological Review 1943 Jul;50(4):370-396. [doi: 10.1037/h0054346]
48. Salgado T, Tavares J, Oliveira T. Drivers of mobile health acceptance and use from the patient perspective: survey study and quantitative model development. JMIR Mhealth Uhealth 2020 Jul 09;8(7):e17588 [FREE Full text] [doi: 10.2196/17588] [Medline: 32673249]
49. Huong DTL, Long TD. Does knowledge matter? The role of m-Health literacy to the acceptance of m-Health applications. JST-UD 2021 Jun 30:42-47. [doi: 10.31130/jst-ud2021-444]
50. Rogers E. Diffusion of innovation. New York, NY: Free Press; 1962.
51. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
52. Rafaeli A, Vilnai-Yavetz I. Emotion as a connection of physical artifacts and organizations. Organization Science 2004 Dec;15(6):671-686. [doi: 10.1287/orsc.1040.0083]
53. McGrenere J, Moore G. Are We All in the Same "Bloat"? 2000 Presented at: Graphics Interface 2000; May 15-17, 2000; Montreal, QC, Canada.
54. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. npj Digit. Med 2019 May 13;2(1):38 [FREE Full text] [doi: 10.1038/s41746-019-0111-3] [Medline: 31304384]
55. Tractinsky N, Zmiri D. Exploring attributes of skins as potential antecedents of emotion in HCI. In: Fishwick PA, editor. Aesthetic Computing . Cambridge, MA: MIT Press; 2006.
56. Tractinsky N, Katz A, Ikar D. What is beautiful is usable. Interacting with Computers 2000 Dec;13(2):127-145. [doi: 10.1016/s0953-5438(00)00031-x]
57. Tuch AN, Roth SP, Hornbæk K, Opwis K, Bargas-Avila JA. Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. Computers in Human Behavior 2012 Sep;28(5):1596-1607. [doi: 10.1016/j.chb.2012.03.024]
58. Thompson DV, Hamilton RW, Rust RT. Feature fatigue: when product capabilities become too much of a good thing. Journal of Marketing Research 2018 Oct 16;42(4):431-442. [doi: 10.1509/jmkr.2005.42.4.431]
59. Montgomery DC, Peck EA, Vining GG, Kowalski SM. Student Solutions Manual to Accompany Introduction to Linear Regression Analysis. Hoboken, NJ: Wiley Blackwell; 1994.
60. Z Score Calculator for 2 Population Proportions. Social Science Statistics. URL: https://www.socscistatistics.com/tests/ztest/default2.aspx [accessed 2022-01-27]
61. Vo V, Auroy L, Sarradon-Eck A. Patients' Perceptions of mHealth Apps: Meta-Ethnographic Review of Qualitative Studies. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: 10.2196/13817] [Medline: 31293246]
62. McKibbin S. Designing web usability: the practice of simplicity. J Direct Data Digit Mark Pract 2001 Jul;3(1):77-78. [doi: 10.1057/palgrave.im.4340116]
63. Karvonen K. The beauty of simplicity. 2000 Presented at: Conference on Universal Usability; November 16-17, 2000; Arlington, VA. [doi: 10.1145/355460.355478]

64. Hilliard ME, Hahn A, Ridge AK, Eakin MN, Riekert KA. User preferences and design recommendations for an mHealth app to promote cystic fibrosis self-management. JMIR Mhealth Uhealth 2014 Oct 24;2(4):e44 [FREE Full text] [doi: 10.2196/mhealth.3599] [Medline: 25344616]
65. Mollerup P. Simplicity. Design Research Quarterly. 2007. URL: https://dl.designresearchsociety.org/cgi/viewcontent.cgi?article=1001&context=design-research-quarterly [accessed 2022-01-27]
66. Mollerup P. Simplicity: A matter of design. Amsterdam, Netherlands: BIS Publishers; 2015.
67. Ramezani Nia M, Shokouhyar S. Analyzing the effects of visual aesthetic of Web pages on users' responses in online retailing using the VisAWI method. JRIM 2020 Sep 17;14(4):357-389. [doi: 10.1108/jrim-11-2018-0147]
68. Li Y, Yeh Y. Increasing trust in mobile commerce through design aesthetics. Computers in Human Behavior 2010 Jul;26(4):673-684. [doi: 10.1016/j.chb.2010.01.004]
69. Hassenzahl M. The Thing and I (Summer of '17 Remix). In: Blythe M, Monk A, editors. Funology 2. Human–Computer Interaction Series. Cham, Switzerland: Springer; 2018:17-31.

Abbreviations

AI: artificial intelligence
mHealth: mobile health
NCD: noncommunicable disease
UI: user interface

Edited by A Kushniruk; submitted 14.03.21; peer-reviewed by D Pförringer, E Baker, E Chukwu; comments to author 04.07.21; revised version received 02.08.21; accepted 02.10.21; published 02.03.22.

Please cite as:
Kela N, Eytam E, Katz A
Supporting Management of Noncommunicable Diseases With Mobile Health (mHealth) Apps: Experimental Study
JMIR Hum Factors 2022;9(1):e28697
URL: https://humanfactors.jmir.org/2022/1/e28697
doi:10.2196/28697
PMID:35234653

Original Paper

Designing Digital Health Technology to Support Patients Before and After Bariatric Surgery: Qualitative Study Exploring Patient Desires, Suggestions, and Reflections to Support Lifestyle Behavior Change

Anna Robinson[1,2], MPharm, PG Clin Dip (Adv); Andrew Husband[1,2], PhD; Robert Slight[2,3], PhD; Sarah P Slight[1,2], PhD

[1]School of Pharmacy, Newcastle University, Newcastle upon Tyne, United Kingdom

[2]Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, United Kingdom

[3]Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom

Corresponding Author:
Sarah P Slight, PhD
School of Pharmacy
Newcastle University
King George VI Building, Newcastle University
Newcastle upon Tyne, NE1 7RU
United Kingdom
Phone: 44 191 208 6000
Email: Sarah.slight@newcastle.ac.uk

Abstract

Background: A patient's capability, motivation, and opportunity to change their lifestyle are determinants of successful outcomes following bariatric surgery. Lifestyle changes before and after surgery, including improved dietary intake and physical activity levels, have been associated with greater postsurgical weight loss and improved long-term health. Integrating patient-centered digital technologies within the bariatric surgical pathway could form part of an innovative strategy to promote and sustain healthier behaviors, and provide holistic patient support, to improve surgical success. Previous research focused on implementing digital technologies and measuring effectiveness in surgical cohorts. However, there is limited work concerning the desires, suggestions, and reflections of patients undergoing bariatric surgery. This qualitative investigation explores patients' perspectives on technology features that would support behavior changes during the pre- and postoperative periods, to potentially maintain long-term healthy lifestyles following surgery.

Objective: This study aims to understand how digital technologies can be used to support patient care during the perioperative journey to improve weight loss outcomes and surgical success, focusing on *what* patients want from digital technologies, *how* they want to use them, and *when* they would be of most benefit during their surgical journey.

Methods: Patients attending bariatric surgery clinics in one hospital in the North of England were invited to participate. Semistructured interviews were conducted with purposively sampled pre- and postoperative patients to discuss lifestyle changes and the use of digital technologies to complement their care. The interviews were audio recorded and transcribed verbatim. Reflexive thematic analysis enabled the development of themes from the data. Ethical approval was obtained from the National Health Service Health Research Authority.

Results: A total of 20 patients were interviewed (preoperative phase: 40% (8/20); postoperative phase: 60% (12/20). A total of 4 overarching themes were developed and related to the optimization of technology functionality. These centered on providing tailored content and support; facilitating self-monitoring and goal setting; delivering information in an accessible, trusted, and usable manner; and meeting patient information-seeking and engagement needs during the surgical pathway. Functionalities that delivered personalized feedback and postoperative follow-up were considered beneficial. Individualized goal setting functionality could support a generation of digitally engaged patients with bariatric conditions as working toward achievable targets was deemed an effective strategy for motivating behavior change. The creation of digital *package of care* checklists between patients and clinicians was a novel finding from this study.

Conclusions: Perceptions of patients undergoing bariatric surgery validated the integration of digital technologies within the surgical pathway, offering enhanced connectedness and support. Recommendations are made relating to the design, content, and functionality of digital interventions to best address the needs of this cohort. These findings have the potential to influence the co-design and integration of person-centered, perioperative technologies.

(JMIR Hum Factors 2022;9(1):e29782) doi:10.2196/29782

KEYWORDS

digital technology; eHealth; mHealth; bariatric surgery; behavior change; qualitative research; co-design; perioperative care; mobile phone

Introduction

Background

Obesity is a growing global pandemic [1-3]. Weight loss surgery (bariatric surgery) is regarded as the most effective method for long-term weight loss [4]. Despite an increase in the number of bariatric procedures over the past few years, recent literature has suggested that surgery is still an underused treatment option, with the number of American adults choosing surgery being approximately 1% [5,6]. Despite the promising weight loss outcomes following surgery, patients can experience challenges beyond the procedure itself in their bid for surgical *success* [7]. These include facing social pressures and stigma related to surgery [8]; psychological impacts, including negative body image and depression [9]; and adjusting to postoperative lifestyle recommendations to reduce weight regain [10].

A patient's capability, motivation, and opportunity to change their lifestyle are significant determinants of successful outcomes following bariatric surgery [11,12]. Healthier lifestyle changes before and after surgery, including improved dietary intake and physical activity levels, have been shown to contribute to greater postsurgical weight loss [13,14], maintenance of weight loss [15], and better overall long-term health [16]. However, previous literature has demonstrated the various challenges that clinicians may face when supporting changes of this nature for this surgical patient cohort, particularly on a long-term basis [17,18]. Attendance at postsurgical follow-up care [19,20], engagement with behavioral appointments and support groups [21,22], and the impacts of travel distance to clinic appointments are some of the previously examined factors associated with poorer outcomes following surgery [17,23]. Digital technologies may pose as a promising alternative avenue for the provision of surgical patient support, which could be offered remotely, without the need for in-person attendance [24]. In particular, digital technologies offer an ability to provide scalable support which may prove useful on a wide scale [24,25]. Currently, little is known about the optimal way to design, deliver, and implement digital health technologies for this unique surgical patient cohort; this study seeks to provide further insights and has adopted a patient-informed and patient-centered approach to do so [14,26].

Digital health technologies (such as mobile phone apps, tailored web platforms, and wearable activity trackers) provide promising opportunities for connected patient care. They provide education and information in an easily accessible and patient-friendly manner [25,27,28] and encourage patients to become active participants in their own care [29,30]. Studies have acknowledged patients' receptiveness toward using digital technologies to complement the care pathways of other surgical procedures, which has resulted in successful behavior change, improved recovery time, and reduced length of stay in hospital [31-33]. In the bariatric surgery literature, recent studies have reported how telemedicine and digitally supported care have been well received by patients [34] and have potentially improved postoperative clinic attendance and patient engagement with surgical care [25,35,36]. Using digital technologies within the bariatric surgical pathway, both pre- and postoperatively, could form part of a remote strategy to deliver support and behavior change advice to patients.

Existing literature has suggested that collaborative approaches in medicine, between patients and clinicians, can result in improved patient engagement, trust, and satisfaction, and improve intended health-related outcomes [37-39]. Cocreation and user-centered, experience-based co-design approaches are being researched and implemented in other areas of health care, with the goal of improving patient-focused care [40,41]. Many studies have focused on implementing digital technologies and measuring their effectiveness in various medical and surgical cohorts [42,43]. A recent study by Korpershoek et al [44] using user-centered design approaches supported patient self-management of chronic obstructive pulmonary disease and that by Solem et al [45] designed and developed an electronic health pain management intervention for those affected by chronic pain. Similarly, a recent study by Paton et al [46] demonstrated how predictive human–computer interaction modeling could be integrated into user-centered design approaches to improve health intervention usability and safety. However, there is a paucity of patient-centered research specifically concerning the desires, suggestions, and reflections of patients undergoing bariatric surgery. This warrants further investigation to develop useful and effective digital support strategies for this patient population, with user-centered design being one possible strategy to adopt to understand how patients undergoing bariatric surgery *want* to be supported.

Objectives

This qualitative study aims to understand how digital technologies could be used to better support patients across the wider perioperative pathway, covering pre- and postoperative time points, with the overall rationale of improving weight loss outcomes and, therefore, surgical success. Specifically, our key research questions were as follows: *What* do patients want from digital health technologies, *How* do they want to use them, and *When* would they be of most benefit during their surgical journey?

Methods

Participant Recruitment and Sampling

According to the Enhancing the QUAlity and Transparency Of health Research guidelines, the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was followed for this study (Multimedia Appendix 1) [47]. Patients attending bariatric surgery clinics within a large teaching hospital in the North of England were invited to participate in this study. This included both pre- and postoperative patients who attended their outpatient appointments, as their experiences and perspectives on using digital health technologies may differ. No previous relationship was established between the researcher and participants before study commencement or recruitment. All participants were provided with an information sheet detailing the purpose and aims of the study during their appointment with the surgeon. Written informed consent was obtained before conducting the interviews. To be included in the study, patients had to be aged >18 years; recently undergone (within the last 2 years, as per the 2-year National Health Service [NHS] bariatric surgery follow-up guidelines) or planned to undergo (ie, those who are under the care of the multidisciplinary team and are awaiting a surgery date) bariatric surgery at the specific hospital trust [48]; medically stable (not affected by an acute decline in health away from baseline); and able to participate in an interview, communicate in English, and have the capacity to consent to taking part in the study. It was deemed important that participants with a range of experiences of and opinions on digital technologies were included in this study to showcase representative views reflective of those of typical patient cohorts; thus, there was no specification to the level of current or previous digital technology use to take part in this study. However, details of the frequency of technology use were collected to aid in the interpretation of the results (Table 1). Purposive sampling techniques were used to recruit a wide and representative sample of patients undergoing bariatric surgery within the region. This also meant that the sample of participants included in this study represented a mixture of ages, men and women, and included participants who were at various stages within their pre- and postoperative journeys (ranging from 1 week before surgery to 24 months after surgery).

Table 1. Participant characteristics (N=20).

Participant number	Sex	Age (years)	Participant ethnicity (self-reported verbatim from participant interviews)	Surgical procedure	Phase	Time since surgery (exact) or time until surgery (approximate[a])	Reported level of digital technology use[b] at the time of interview
1	Female	29	“English”	Gastric bypass	Postoperative	24 months	Daily
2	Female	55	“White British”	Sleeve gastrectomy	Postoperative	12 months	Daily
3	Female	54	“Pakistani Asian”	Gastric band	Postoperative	18 months	Daily
4	Female	50	“British”	Sleeve gastrectomy	Postoperative	24 months	Daily
5	Male	46	“British”	Undecided	Preoperative	6 weeks	Every other day
6	Female	52	“British”	Gastric bypass	Postoperative	9 months	Daily
7	Female	61	“English”	Gastric bypass	Postoperative	4 months	Never
8	Male	51	“British”	Gastric band	Postoperative	24 months	Daily
9	Female	39	“White British”	Sleeve gastrectomy	Preoperative	2 weeks	Daily
10	Male	40	“Asian”	Gastric bypass	Preoperative	8 weeks	Daily
11	Female	31	“British”	Gastric bypass	Postoperative	24 months	Daily
12	Female	51	“British”	Gastric bypass	Postoperative	24 months	Daily
13	Female	58	“White British”	Gastric bypass	Postoperative	24 months	Daily
14	Female	50	“White British”	Gastric bypass	Preoperative	1 week	Daily
15	Female	59	“English”	Gastric bypass	Postoperative	24 months	Every other day
16	Female	29	“Pakistani”	Gastric bypass	Postoperative	12 months	Daily
17	Male	26	“Asian”	Sleeve gastrectomy	Preoperative	8 weeks	Daily
18	Female	35	“British”	Gastric band	Preoperative	4 weeks	Daily
19	Male	50	“White British”	Undecided	Preoperative	2 weeks	Daily
20	Female	52	“British Indian”	Gastric bypass	Preoperative	4 weeks	Daily

[a]Time until surgery, approximate: given the implications of the COVID-19 pandemic, some surgery dates may have been delayed.

[b]Reported level of digital technology use: reported by participants in response to the question *How often do you use the internet or use apps on a smartphone?*

Semistructured Interview

Between February and March 2020, in-depth semistructured interviews were conducted by a researcher (AR, a female doctoral researcher with experience in qualitative research). All participants chose to be interviewed in the hospital in a confidential surgery clinic room, at a time convenient for them; only the participant and researcher were present. Interviews were conducted until theoretical data saturation was reached, that is, upon author consensus that subsequent interviews yielded no new information. Instead, the authors observed mounting instances of the same codes, as described by Urquhart et al [49], Birks and Mills [50], and Olshansky and de Chesnay [51], and deemed that theoretical data saturation had been achieved. A semistructured interview schedule (topic guide), which formed the basis of all participant interviews, was developed based on 3 pilot interviews, existing studies on digital health technologies in this cohort [27,28,52,53], and systematic reviews of the literature by the research team [24,25]. Participant interviews included questions to elicit spontaneous discussions around their surgical experience, awareness of health and lifestyle behavior change (eg, physical activity, smoking cessation, alcohol intake, and dietary intake), patient physical and psychological support requirements, their perspectives on digital health technologies, and previous technology use (Item 2: Topic guide in Multimedia Appendix 1).

Data Analysis

Semistructured interviews were audio recorded and transcribed verbatim by a researcher (AR). All data were anonymized at the point of transcription. Participants did not provide comment on the transcript or feedback on results. Each interview was transcribed and analyzed before conducting the next interview. Reflexive thematic analysis, as defined by Braun and Clarke [54,55], was performed by 2 researchers (AR and AKH). Transcribing the audio files and reading and rereading the interview transcripts ensured data familiarization. Significant phrases and sections of transcripts were identified and coded with initial descriptive codes; these were then sorted and clustered into common coding patterns, which enabled the development of analytic themes (derived from the data). Working iteratively and reflexively, the themes were continuously reviewed and refined until they were coherent and distinctive [54]. Reflexive analysis was performed through discussion between the 2 researchers (AR and AKH) and, if agreement was not reached, by consensus with the wider team (SPS and RDS). Postinterview field notes enhanced this reflective process. NVivo (version 12; QSR International) was used to assist in the organization of interview data and thematic analysis. The team members were in agreement that data saturation occurred at 20 interviews. When using direct quotes from patients, nonidentifiable pseudonyms were used to ensure confidentiality; for example, participant 1, participant 2, and so on.

Ethical Approval

Ethical approval was obtained from the NHS Health Research Authority and Care Research Wales (reference 19/NE/0318).

Results

Overview

A total of 20 participants were recruited and interviewed as part of this study (there were no refusals to partake, participant dropouts, or repeat interviews). Of these 20 participants, 8 (40%) participants were in the preoperative phase and 12 (60%) were in the postoperative phase of their surgical journey. The characteristics of each participant are presented in Table 1. The average age of participants was 46 (SD 10.63) years, and most of the participants had, or were planning to undergo, a gastric bypass procedure (11/20, 55%). All patient interviews were conducted in person between February and March 2020, before the COVID-19 pandemic and restrictions. All participants chose to be interviewed in a confidential room within the bariatric surgery clinic of the hospital. The average interview duration was 52 (SD 18.5) minutes.

The analysis revealed that participants had particular support needs throughout their perioperative journey before and after bariatric surgery. A total of 4 overarching themes were developed from data related to the *capability* and *functionality* of digital health technologies to provide this support. These concerned the technology's ability to (1) provide surgery-specific content and support; (2) facilitate self-monitoring and goal setting; (3) deliver information in an accessible, trusted, and usable manner; and (4) meet information-seeking and engagement needs at time points before and after undergoing bariatric surgery. We further explored these 4 themes and illustrated perspectives and suggestions with direct interview quotes within this patient-informed piece of work.

Providing Surgery-Specific Content and Support

When asked about *how* digital technologies could best be designed for patients undergoing bariatric surgery, interviewees expressed opinions about what information should be provided, how this information should be tailored, how specific features could be designed, and their visions of what their *ideal* supportive digital intervention would look like.

It was deemed important that the content and support that patients received from the technology were specific to bariatric surgery. A preoperative participant described how "the support packages should be tailored to the people, rather than the procedure," explaining how patients "can lose our hair, end up with excess skin, and need to be on lifelong supplements" and how this is "the kind of stuff" that they need support with throughout the journey of surgery and beyond (participant 14, preoperative phase). Another participant explained how it would have been helpful to know that "after a normal operation you'd be able to eat whatever to build up your energy levels again quite quickly...but you can't do that with bariatric surgery, you physically can't eat things immediately post-surgery," so "you'd need it specifically to advise on the bariatric recovery in that case" (participant 3, postoperative phase). There appeared to be an unmet need related to tailored, educational, and informational support for this cohort.

Regarding the content of the technology, discussions centered on dietary-focused forms of support. Patients' suggestions and desires ranged from the inclusion of "options of what I could have for a snack" (participant 5, preoperative phase) and "something with a meal plan available" (participant 9, preoperative phase), to designing "an app with recipes on it" so patients could "keep coming back to it" for healthier meal options (participant 5, preoperative phase). Patients favored prescriptive approaches (defined as stating what should happen or what someone should do) to content when it came to describing *ideal* technology-enabled support, stating that the intervention should tell them what to do and what to "stick to" (participant 8, postoperative phase). A preoperative patient suggested that the integration of features such as "a list of what you're not allowed to eat anymore" would be most helpful so they could "easily keep away from it (unhealthy foods)" in a bid to "keep on track" with their anticipated weight loss (participant 14, preoperative phase). In a cohort required to change their lifestyle behaviors, even before undergoing surgery, perhaps technologies delivering short-term descriptive support (defined as describing something in a detailed way) would be beneficial. Participants also stated that immediately following surgery, they wished for stricter prescriptive digital support to help them adjust to their new postoperative lifestyle and dietary intake:

> *In the first couple of weeks [following surgery], we need to be told what to do, what exactly to do...like what to eat and what to avoid* [Participant 9, preoperative phase]

Participants considered it important for technology content to also focus on the wider elements of healthy lifestyle behaviors, including increased physical activity and reduced alcohol intake: "If you called it a 'lifestyles package' for after bariatric surgery then you can mention things like diet but also [alcohol] drinking and exercise" (participant 9, preoperative phase). Patients demonstrated awareness that positive behavior changes in these areas also contributed to bariatric surgery success, with a participant specifically discussing how they were "trying to look for better choices – like a better choices app" (participant 5, preoperative phase) to support their journey. Interviewees described how building reminders and prompts into technology could better promote these messages of positive health behaviors. The tone and content of these prompts were perceived to be important, combining monitoring and activity messages with motivational statements. The same participant described how patients should be given control over the technology settings so they could decide on the correct tone for them.

> *I would want something to just give you little reminders – maybe even "have you been weighed this week?" "have you been for a walk?" "don't let yourself slip", things like that. But erm, nothing too forceful...Not the whole powered sort of, gym messages, like "get up fatty!" [laughs]* [Participant 10, preoperative phase]

A postoperative patient reflected that, regardless of the technology delivery method used, "the most important thing is that you're not left alone after the operation...[as] there's so many unknowns [sic]" (participant 11, postoperative phase). Instead, they called for tailored digital support to be on hand throughout the entire surgical journey to provide reassurance to patients both pre- and postoperatively.

Facilitating Self-monitoring and Goal Setting

Both pre- and postoperative participants reflected on the usefulness of self-monitoring and goal-setting functionalities to track their progress throughout the surgical journey. Participants felt it would be useful to self-monitor with "comparison photos" that could be uploaded to an app to "see how much of a difference there has been" (participant 15, postoperative phase). Participants discussed real-time engagement with technologies, remarking the usefulness of inputting daily or weekly weights so that "graphs can track" (participant 7, postoperative phase) and visualize their total weight loss over time. Self-monitoring features were also discussed in association with motivation and emotional investment in the surgical journey, where a participant described how observing "how much [weight] you've lost" can "keep people's spirits up" (participant 15, postoperative phase). Another participant explained how automated messages of "congratulations" were encouraging and "if it calculates your BMI going down as well, I think that would be a really good motivational tool" (participant 7, postoperative phase). Suggestions to incorporate digital self-monitoring features into digital interventions appeared to acknowledge the determination of this cohort in striving for surgical success.

Patients also recognized how self-monitoring could encourage and *push* them to undertake positive health behaviors related to their physical activity to support their postoperative weight loss. A participant described how wearable technology enticed them "into doing more steps or exercise" (participant 1, postoperative phase) and another referred to using gamification features with different *levels* of increased difficulty for them to work through. This participant suggested that increasing the step count targets on a monthly basis would challenge them to continue with regular walking and that achieving the target meant they were encouraged to walk further for the next month. Progressing through these physical activity–based milestones was seen to encourage engagement with their physical rehabilitation and provide underlying reassurance of staying *on track* with their recovery.

> *I'd want [the physical activity challenges] to have different levels too - like the first month, the second month, unlocking the next bit...Then it's all there for you and you can keep going back and checking on the app...I can know I'm on track then* [Participant 14, preoperative phase]

A participant described a common postoperative pitfall of getting "so hung up on what we're eating and whether it's right or wrong" (participant 11, postoperative phase). Instead, they recognized the benefit of technological features that enable the setting of "daily goals about exercise" to "give us something else to think about...and work towards" (participant 11, postoperative phase), while achieving their vision of optimal postoperative weight loss. The same participant reflected on how goal setting would have widened their personal knowledge of "what to do after" surgery, meaning they were able to

"recover better" (participant 11, postoperative phase) and more successfully. Another participant drew on their personal experiences of using the "NHS Patient Access app" (participant 7, postoperative phase), which is freely available for all patients registered with a general practitioner (physician) in the United Kingdom. The app can be used to view primary care health records, order repeat prescriptions, and get health advice on medical conditions and treatments [56]. This participant suggested that there be inclusion of specialist-bariatric advice within the app, where "the full app [could be linked] to your NHS number so it's all personalized advice available" (participant 7, postoperative phase). The participant also suggested useful additions to the NHS app, where the home screen could include "tabs at the bottom for specific stuff...like graphs to track [your progress]" (participant 7, postoperative phase).

Participants also discussed the value of shared access to their self-monitored data, where members of the multidisciplinary surgical team were able to track their progress. They remarked that in-built 2-way monitoring features could increase their personal sense of motivation and accountability to "break those [bad] habits" participant 10, preoperative phase), especially knowing that someone else was "keeping an eye" (participant 11, postoperative phase). Another participant felt that shared monitoring could act as a reassurance mechanism for patients, in which they were not being left to "fend for themselves" (participant 4, postoperative phase) in the run-up to surgery or as soon as the surgery was over. A sense of shared responsibility for the success of surgeries was discussed when considering professional-led health care monitoring. A participant supported the inclusion of shared monitoring capabilities so that both patients and health care professionals can "notice if they're slipping" (participant 16, postoperative phase) off the postsurgical diet, implying that patients alone may not be able to recognize bad habits reforming.

Delivering Information in an Accessible, Trusted, and Usable Manner

All participants offered suggestions on technology delivery methods and how they would like the intervention to be available to them, including via phone-based apps, web-based forums, and the use of social media platforms such as Facebook. Most participants discussed that their preferred delivery method would be accessible through their smartphone via an app, with a patient explaining "practically everyone knows how to use a phone for stuff now. Everything's on it...So, if you could put an app on there, I reckon that's the best way" (participant 15, postoperative phase). Other participants also reported how frequently they used their phones and how people rarely "go anywhere without it," offering the potential for ongoing engagement even "if I'm out for the day or away on holidays or whatever, I can still log in" (participant 14, preoperative phase) to use it. Many interviewees desired a delivery system that was "nice and clear" (participant 3, postoperative phase), with one remarking they did not want another "dry or crisp NHS website," instead preferring a "modernized" (participant 4, postoperative phase) app or discussion page.

As an alternative delivery method, some participants reported being members of bariatric surgery groups on Facebook. A few participants reported social media and Facebook to be an acceptable delivery format, offering familiarity and reassurance: 'I use Facebook all the time...it's amazing' (participant 9, preoperative phase). However, participants also questioned the reliability of information posted on Facebook, describing it as "obviously everyone's own experiences, but it might not necessarily be the safest" (participant 11, postoperative phase). A participant described how some of the posts they had read were "full of nonsense," and therefore, they got rid of their account. In their view, "an app would be better" as they "would probably trust it [the content] more than Facebook" (participant 5, preoperative phase). Furthermore, another drawback of Facebook was how one "need[ed] to scroll back to find the information," whereas an app could contain "a specific folder or tab so you could go back to it [information]" (participant 9, preoperative phase). Other participants described their positive experiences of *closed* groups with smaller numbers of individuals. A female patient discussed a private WhatsApp group which contained 5 other postoperative patients and felt that the "how are you all doing? messages" (participant 4, postoperative phase) were helpfully shared among themselves. This indicates that some postoperative patients may find it helpful to surround themselves with smaller groups of like-minded individuals when seeking trusted information.

Many participants highlighted how any information needs to be quick and easy to locate, with one suggesting it should be kept "all together in one place" (participant 9, preoperative phase) and another describing how "that way you can keep coming back to the information any time you wanted to, rather than looking for the leaflets they gave us" (participant 5, preoperative phase). Another described organizing the information with "tabs at the bottom [of the screen] for specific stuff" like "appointments for follow ups" (participant 7, postoperative phase).

Previous technology use was considered along with accessibility and information provision. A participant described "a usable manner" as something that depends "on your character. I'm not very techno-loving or anything, but I'd give it a go [laughs]" (participant 6, postoperative phase). Some participants discussed usability from the perspective of others, particularly the older family members. A interviewee considered her mother aged 63 years, describing how "she can use Google now, but it's took a long time to get her to do that [sic]. But then again, my husband's Dad, he's 73 and he would definitely use digital stuff." Interestingly, she also appreciated that usability "is a bit dependent on the person too, not just their age" (participant 9, preoperative phase). Some interviewees viewed usability in the same context as familiarity and referred to strategies to overcome this through patient education.

Another participant offered suggestions of how to design the technology so that users of all literary abilities could engage, through the use of *happy* or *sad* faces, or colors, for instance:

> *I've met a lot of people that can't read or write...you could do happy face, sad face, whatever...Or amber*

> *color for not advisable, red for bad or danger, green for good* [Participant 12, postoperative phase]

Meeting Patient Information-Seeking and Engagement Needs at Time Points Before and After Surgery

With regard to using a form of digital technology for support, participants shared varying opinions about *when* it would be of most benefit to them during the perioperative period. This benefit appeared to relate to (1) the timing of intervention *implementation* (eg, implementing the technology to enable preoperative information seeking) and (2) the timing of desired *engagement* with technologies (eg, the value of interventions that offered functions that spanned short term and long term to meet patient needs).

When considering their implementation within the surgical journey, participants believed preoperative digital interventions would be useful to acquire knowledge about their upcoming surgery "it's an operation at the end of the day and you're changing your insides so I think it's important to fully know [about] it" (participant 10, preoperative phase). Participants considered this preoperative knowledge-forming period vital for both their physical and mental preparedness. After struggling with their own surgical outcomes, a participant suggested preoperative digital support specifically relating to the psychological preparation of surgery. They discussed how preoperative interventions could better educate patients and meet information-seeking needs and manage postoperative weight loss expectations:

> *If something could teach me like how to expect, what to expect after [the surgery], it might have helped..."cause I thought the weight loss would be much faster and I look no different now, which has affected my mental health.*" [Participant 3, postoperative phase]

Similar thoughts were raised by other participants, with one explaining how it "would be really useful to have a map or plan to know what's going to happen, and when, so we know it's a full process for us to refer to and not panic" (participant 4, postoperative phase). Another suggested designing "a checklist...like all part of your own bariatric package" where you could "tick off each bit" when it was achieved (participant 3, postoperative phase). Patients may benefit from seeing the phases of the journey and understanding what was going to happen next:

> *At least you could know what to expect, what is coming either before or after the procedure, and what to do.* [Participant 9, preoperative phase]

Interviewees recognized the value of information seeking in the initial, short-term, postoperative period "cause, say you were standing in the supermarket and you thought 'oh I could really fancy that, but I don't know if I'm allowed it' then you'd be able to look it up and see if you can have it or not. That would be really practical and handy" (participant 14, preoperative phase). Interviewees recognized that engagement with technologies would likely be higher in the initial postoperative period "once you've had it [surgery], you're in it, and probably will need the information there and then..." (participant 10, preoperative phase), but that each participant's engagement needs will change, further along their postsurgical surgery they are. Participants also considered the role that technologies could play in terms of long-term ongoing support, where the ability to engage with an intervention, when needed, was deemed important:

> *It might be something where it [intervention usage] tails off a bit, once you start getting the hang of things, what to eat, how much you can tolerate and stuff. But also, if anything happened and I wanted to ask questions, then I picture being able to use it as and when.* [Participant 14, preoperative phase]

Two participants (one in the preoperative phase and another in the postoperative phase) acknowledged that technologies could play a role in complementing current practices to improve patient support between annual follow-up appointments. A postoperative participant explained that "once you got a few months in it was more 'well, I'll see you in 12 months unless you have problems' and that's not supportive enough" (participant 11, postoperative phase). They believed there to be benefit from continued technology-enabled engagement throughout this time, specifically linking with a health care professional for advice: "if I'd had more contact with the dietician, digitally, I could maybe have stayed on track better" (participant 11, postoperative phase). Recurring messages of prescriptive and descriptive approaches, in which postoperative participants appear to cede complete control over their journey and outcomes, perhaps demonstrates a lack of belief that they can make and sustain positive behavior changes on their own. A preoperative participant perceived the value of ongoing support from technologies in a more self-determined manner: "I want to make sure I get it [dietary intake] right. I want to avoid any complications and give myself the best chance of success" (participant 5, preoperative phase). They went on to describe their ideal technology-enabled support system, combining technology alongside face-to-face appointments, stating: "I think using tech and still having the [face-to-face] appointments will give me as much support as I need" (participant 5, preoperative phase).

Of all the participants interviewed, only one recommended implementing an intervention that spanned both the pre- and postoperative periods. This patient was in the 2-year postsurgery phase and their views combined those of the pre- and postoperative patients discussed in the previous sections. They described how supportive *boosts* from the technology, continued on a long-term basis, could help to promote positive behaviors:

> *From the minute you decide to go through with it [surgery], you probably would benefit from having something there just for peace of mind...definitely [implementing] from the start, but also so they can keep using it after [surgery] too for those little boosts and support.* [Participant 15, postoperative phase]

Discussion

Principal Findings

This patient-informed study identified the desires, suggestions, and reflections of bariatric patients in the context of using digital health technologies as support tools during surgery. By collecting both pre- and postoperative patient perspectives, we highlighted *how* digital support strategies could be delivered, *what* content is perceived as useful, and *when* technologies could be implemented within the current NHS bariatric surgery pathway. Our findings discussed 4 key themes related to technology functionality and capability that enable better tailored and targeted digital health technologies for bariatric surgical patients.

Limitations

Our results have important implications for the design, delivery, usability, and implementation of digital technologies for patients undergoing bariatric surgery. Uniquely, our findings collate participant desires, suggestions, and reflections concerning digital technology use across the entire bariatric perioperative pathway. This study is one of the first to incorporate pre- and postoperative participants, building evidence on the optimization of technology-based support to span the perioperative journey when undergoing bariatric surgery. We acknowledge that there were some limitations to this study. First, the research predominantly focused on a small sample of patients in the North of England. Second, as is common with bariatric surgery, this sample included more female participants than male participants. In addition, we did not assess or sample participants according to their socioeconomic status; it is possible that participants of different socioeconomic classes may have varied experiences with technologies, and our results should, therefore, be interpreted with this in mind. Participants included in this study were purposively sampled from attendees at bariatric surgical clinics (including preoperative assessments and postoperative follow-up appointments); thus, the results do not include patients who were under hospital care but were noncompliant with appointment attendance. Further research that specifically focuses on the experiences and perceptions of participants from ethnic minority communities undergoing bariatric surgery is needed, given that 75% (15/20) of this sample self-reported British or White British ethnicity. Finally, our study also focused solely on the desires, suggestions, and reflections of bariatric surgical patients; thus, the results may not be generalizable to other elective surgical procedures. Future studies may wish to deepen the insights gained from this study to more closely consider patient journey and changing mindsets from pre- to postsurgery phases, which may affect the rates of patient engagement with technologies.

Comparison With Previous Work

Study participants described a range of potential technological suggestions to meet their pre- and postoperative needs. Patients discussed how digital health technologies could be implemented to enable access to specialist information specific to bariatric surgery, located in an easily accessible place. They demonstrated preferences for digital interventions that incorporated content specific to bariatric surgery rather than being focused on generalized nonsurgical weight loss. Comparable with findings in wider digital health literature, the patients in this study also highlighted the benefits of functionalities that offer support on an individualized basis, such as enabling the provision of individualized feedback and personalized reviews on postoperative progress [57]. Personalization of feedback has previously been associated with positive health behavior changes and increased patient engagement with care [58-60]. A participant suggested connecting technologies to health system identifiers, such as an individual's NHS number, to support the delivery of personalized care.

In line with this study, perspectives of becoming *digitally engaged patients* were discussed by many participants [61]. For this cohort, the focus of their engagement centered on the monitoring of postoperative progress, primarily the ability to track surgically induced weight loss. Previously, web-based health technologies with monitoring capabilities have been credited as transformers of health care by supporting engaged self-care and promoting positive health behaviors [62]. In addition to individualized feedback, the potential for individualized goal setting may further support the generation of digitally engaged patients with bariatric conditions. Working toward achievable targets has been deemed an effective strategy to successfully motivate behavior change [63]. Wider literature echoes that individualized goal setting has demonstrated improvements in sedentary behavior [64,65], personalized feedback and messages of encouragement have provided patients with cancer, a sense of accomplishment [66], and visual tracking of physical activity (eg, daily step counts) has been reported as motivational [65,67]. Perhaps the same approach could be used for patients undergoing bariatric surgery, with a focus on achievable targets of weight loss, combined with dietary intake and physical activity. Uniquely, a participant reflected on gamification when designing technologies (in game format) to support staged surgical recovery. A study focusing on increased physical activity to aid recovery following cancer surgery expressed similar findings; these authors also identified that personalized difficulty settings in the *game* boosted patient satisfaction and engagement with the intervention [68].

Numerous participants referred to the surgical journey as a process, suggesting that it may benefit from technology-enabled checklists to create a *package of care* between patients and clinicians—a novel finding from this study. Patients envisaged this to be of particular use in the early postoperative period, enabling better control over their recovery and diet and a better understanding of their follow-up care. References were made to design helpful prompts for patients. This echoed previous findings where the tone and delivery of these prompts or messages were deemed crucial in motivating sustained positive health behaviors in patients with cancer [67,69,70].

There appears to be value in implementing technologies both preoperatively and postoperatively. Echoing participant reflections in this study, preoperative interventions have previously been linked with promoting positive behavior change culture [27,28,71]. This is closely linked with theories of surgical teachable moments, arguing that patients are highly susceptible and motivated to change following the initial decision to undergo surgery [72,73]. Highlighting the

perspectives of the participants in this study, digital health technologies may present a promising opportunity to prepare patients before surgery and provide continued support between routine postoperative follow-up appointments.

The timing of engagement with technologies appeared to be individualized. The results from this study suggest that, in addition to using technology on a regular basis for personalized prompts and messages, some participants highlighted a desire to engage with the technology on an ongoing basis. The benefit of being able to engage *when required* seems logical, particularly for a patient cohort with changeable postoperative needs over time. The participants in this study also considered that intervention use and engagement rates would likely be higher soon after surgery but reduce over time once they better adjusted to their life after surgery. The dichotomy concerning intervention timings revealed in this study draws attention to the importance of finding optimal *engagement balance* with any digital health technologies implemented for patients. Currently, there is insufficient evidence to state the optimal initiation point and ongoing engagement points of digital technologies within the bariatric surgical pathway, an area that future studies may explore further.

Participants raised contrasting views that suggested a fine balance existed between them accepting and abdicating responsibility over their recovery and subsequent surgical *success*. Prescriptive and descriptive approaches to technology content were desired by some who wanted the technology to provide them with regulated and specific advice, such as directed postoperative meal plans. However, previous studies have noted this approach to have questionable success when it comes to motivating and sustaining behavior change [74]; instead, the authors have cited the importance of empowered patient–health provider strategies [75,76]. Self-determination Theory (SDT) provides a theoretical framework to understand participant motivations and behaviors [77]. When SDT was applied to other health behavior contexts (such as programs for smoking cessation [78] and weight loss [79]), findings suggested that the more autonomously motivated participants were, the more successfully they implemented behavior change. The information-motivation-behavioral skills model of health behavior has been widely used in medical research [80-82] to understand and improve patient health behaviors and increase the efficacy and effectiveness of behavioral interventions. This model states that educational information (which could be prescriptive in nature, as desired by this cohort) is a prerequisite to successfully enact a change in health behaviors [83]. Both the SDT and information-motivation-behavioral skill models propose that patients who are well educated and informed, with higher levels of independent motivation and acceptance of responsibility, are more likely to enact and maintain health-related behaviors. In the context of this study, the desire for prescriptive and descriptive approaches to technology content is not necessarily at odds with the need for interventions that boost patient motivation; both approaches may be regarded as requirements for supporting successful patient weight loss, both in the short term and long term.

Technology-enabled monitoring has also been recognized to boost autonomous motivational levels [77]; however, long-term monitoring by health care professionals as desired by the patients may be considered unsustainable. Monitoring opportunities and timescales should be considered when it comes to digital technology design and functionality to support and motivate patients during their surgical journey. Given its value as a source of potential accountability and motivation for self-monitoring and social support benefits, digitally-enabled peer networking within the bariatric surgical journey should be considered as an area for future research, in particular, *how* and *when* digital health technologies could support with, and facilitate, this [25]. Future research should focus on the motivational role of digital technologies when providing support to patients facing challenges within the bariatric surgical pathway, such as regaining weight.

Similar to previous digital health research, themes of usability were discussed by participants, particularly regarding their existing familiarity versus unfamiliarity with technologies [84]. Reflections from the perspective of older relatives highlighted that digital literacy and generational bias may still be a challenge to overcome when considering the implementation of health technologies [67,85,86]. Although technologies are now implemented more readily within health care, some patients may still prefer face-to-face encounters with clinicians rather than web-based ones [61]. We should be mindful of acknowledging this and, as suggested by the participants, work to complement technological integration with educational support materials.

Conclusions

Perceptions of patients undergoing bariatric surgery validate the integration of digital health technologies within the surgical care pathway, offering enhanced connectedness and support. The findings from this study have the potential to influence the design and targeting of future digital technologies to best support bariatric surgical patients. To achieve surgical success, digital strategies should consider the incorporation of specialist information tailored to the bariatric surgery cohort and the implementation of self-monitoring and goal-setting functionalities at various time points within the bariatric surgical pathway. Further, to address the specific unmet support needs of this patient cohort, digital health technologies should enable the provision of a *package of care* to offer long-term lifestyle support.

Acknowledgments

The authors would like to thank all the participants who participated in the interviews and offered their perspectives to make this research possible. In addition, they would also like to thank Dr Neil Jennings and Dr Kamal Mahawar (bariatric surgeons), Deila Pearce (staff nurse), and the bariatric surgical multidisciplinary team for their support with data collection. AR is in receipt of

the Dr WE Harker PhD Studentship from Newcastle University, which covers their doctoral candidature and relevant publication costs.

Authors' Contributions

All the authors were responsible for the conception and design of the study. AR performed data collection. AR and AH performed data analysis and liaised with RS and SPS, as required. AR led the writing of this manuscript as part of their PhD doctoral candidature, with all the coauthors commenting on various drafts. All the authors have read and approved the final manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Qualitative research checklist and interview topic guide.
[DOCX File , 43 KB - humanfactors_v9i1e29782_app1.docx]

References

1. Kelly T, Yang W, Chen C, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. Int J Obes (Lond) 2008 Sep 8;32(9):1431-1437. [doi: 10.1038/ijo.2008.102] [Medline: 18607383]
2. Hruby A, Hu FB. The epidemiology of obesity: a big picture. Pharmacoeconomics 2015 Jul;33(7):673-689 [FREE Full text] [doi: 10.1007/s40273-014-0243-x] [Medline: 25471927]
3. Agha M, Agha R. The rising prevalence of obesity: part A: impact on public health. Int J Surg Oncol (N Y) 2017 Aug;2(7):e17 [FREE Full text] [doi: 10.1097/IJ9.0000000000000017] [Medline: 29177227]
4. Panteliou E, Miras AD. What is the role of bariatric surgery in the management of obesity? Climacteric 2017 Apr 04;20(2):97-102. [doi: 10.1080/13697137.2017.1262638] [Medline: 28051892]
5. Campos G, Khoraki J, Browning M, Pessoa B, Mazzini G, Wolfe L. Changes in utilization of bariatric surgery in the United States from 1993 to 2016. Ann Surgery 2020 Feb 2;271(2):201-209. [doi: 10.1097/sla.0000000000003554]
6. English W, DeMaria EJ, Brethauer S, Mattar S, Rosenthal R, Morton J. American society for metabolic and bariatric surgery estimation of metabolic and bariatric procedures performed in the United States in 2016. Surg Obes Relat Dis 2018 Mar;14(3):259-263. [doi: 10.1016/j.soard.2017.12.013] [Medline: 29370995]
7. McGrice M, Don Paul K. Interventions to improve long-term weight loss in patients following bariatric surgery: challenges and solutions. Diabetes Metab Syndr Obes 2015;8:263-274 [FREE Full text] [doi: 10.2147/DMSO.S57054] [Medline: 26150731]
8. Han S, Agostini G, Brewis AA, Wutich A. Avoiding exercise mediates the effects of internalized and experienced weight stigma on physical activity in the years following bariatric surgery. BMC Obes 2018 Jul 2;5(1):18-19 [FREE Full text] [doi: 10.1186/s40608-018-0195-3] [Medline: 29988619]
9. Ivezaj V, Grilo CM. The complexity of body image following bariatric surgery: a systematic review of the literature. Obes Rev 2018 Aug;19(8):1116-1140 [FREE Full text] [doi: 10.1111/obr.12685] [Medline: 29900655]
10. Bradley LE, Forman EM, Kerrigan SG, Goldstein SP, Butryn ML, Thomas JG, et al. Project HELP: a remotely delivered behavioral intervention for weight regain after bariatric surgery. Obes Surg 2017 Mar;27(3):586-598. [doi: 10.1007/s11695-016-2337-3] [Medline: 27586525]
11. Thomas JG, Bond DS, Sarwer DB, Wing RR. Technology for behavioral assessment and intervention in bariatric surgery. Surg Obes Relat Dis 2011;7(4):548-557 [FREE Full text] [doi: 10.1016/j.soard.2011.02.009] [Medline: 21514246]
12. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
13. Courcoulas AP, Christian NJ, Belle SH, Berk PD, Flum DR, Garcia L, Longitudinal Assessment of Bariatric Surgery (LABS) Consortium. Weight change and health outcomes at 3 years after bariatric surgery among individuals with severe obesity. J Am Med Assoc 2013 Dec 11;310(22):2416-2425 [FREE Full text] [doi: 10.1001/jama.2013.280928] [Medline: 24189773]
14. Stewart F, Avenell A. Behavioural interventions for severe obesity before and/or after bariatric surgery: a systematic review and meta-analysis. Obes Surg 2016 Jun;26(6):1203-1214. [doi: 10.1007/s11695-015-1873-6] [Medline: 26342482]
15. Sarwer DB, Dilks RJ, West-Smith L. Dietary intake and eating behavior after bariatric surgery: threats to weight loss maintenance and strategies for success. Surg Obes Relat Dis 2011 Sep;7(5):644-651 [FREE Full text] [doi: 10.1016/j.soard.2011.06.016] [Medline: 21962227]
16. Sjöström L, Peltonen M, Jacobson P, Ahlin S, Andersson-Assarsson J, Anveden A, et al. Association of bariatric surgery with long-term remission of type 2 diabetes and with microvascular and macrovascular complications. J Am Med Assoc 2014 Jun 11;311(22):2297-2304. [doi: 10.1001/jama.2014.5988] [Medline: 24915261]

17. McVay MA, Friedman KE, Applegate KL, Portenier DD. Patient predictors of follow-up care attendance in Roux-en-Y gastric bypass patients. Surg Obes Relat Dis 2013;9(6):956-962 [FREE Full text] [doi: 10.1016/j.soard.2012.11.005] [Medline: 23295164]
18. Shilton H, Gao Y, Nerlekar N, Evennett N, Ram R, Beban G. Pre-operative bariatric clinic attendance is a predictor of post-operative clinic attendance and weight loss outcomes. Obes Surg 2019 Jul;29(7):2270-2275. [doi: 10.1007/s11695-019-03843-2] [Medline: 30903430]
19. Odom J, Zalesin KC, Washington TL, Miller WW, Hakmeh B, Zaremba DL, et al. Behavioral predictors of weight regain after bariatric surgery. Obes Surg 2010 Mar;20(3):349-356. [doi: 10.1007/s11695-009-9895-6] [Medline: 19554382]
20. Gould JC, Beverstein G, Reinhardt S, Garren MJ. Impact of routine and long-term follow-up on weight loss after laparoscopic gastric bypass. Surg Obes Relat Dis 2007;3(6):627-630. [doi: 10.1016/j.soard.2007.07.005] [Medline: 17950045]
21. Papalazarou A, Yannakoulia M, Kavouras SA, Komesidou V, Dimitriadis G, Papakonstantinou A, et al. Lifestyle intervention favorably affects weight loss and maintenance following obesity surgery. Obesity (Silver Spring) 2010 Jul;18(7):1348-1353. [doi: 10.1038/oby.2009.346] [Medline: 19834466]
22. Gallé F, Marte G, Cirella A, Di Dio M, Miele A, Ricchiuti R, et al. An exercise-based educational and motivational intervention after surgery can improve behaviors, physical fitness and quality of life in bariatric patients. PLoS One 2020;15(10):e0241336 [FREE Full text] [doi: 10.1371/journal.pone.0241336] [Medline: 33119658]
23. Khorgami Z, Zhang C, Messiah SE, de la Cruz-Muñoz N. Predictors of postoperative aftercare attrition among gastric bypass patients. Bariatr Surg Pract Patient Care 2015 Jun 01;10(2):79-83 [FREE Full text] [doi: 10.1089/bari.2014.0053] [Medline: 26155436]
24. Robinson A, Husband A, Slight R, Slight S. Digital technology to support lifestyle and health behaviour changes in surgical patients: systematic review. BJS Open 2021 Mar 05;5(2):2020 [FREE Full text] [doi: 10.1093/bjsopen/zraa009] [Medline: 33688953]
25. Robinson A, Husband AK, Slight RD, Slight SP. Digital support for patients undergoing bariatric surgery: narrative review of the roles and challenges of online forums. JMIR Perioper Med 2020 Jul 15;3(2):e17230 [FREE Full text] [doi: 10.2196/17230] [Medline: 33393926]
26. Eaton L, Walsh C, Magnuson T, Schweitzer M, Lidor A, Nguyen H, et al. On-line bariatric surgery information session as effective as in-person information session. Surg Obes Relat Dis 2012 Mar;8(2):225-229 [FREE Full text] [doi: 10.1016/j.soard.2011.10.015] [Medline: 22178563]
27. Lemanu DP, Singh PP, Shao RY, Pollock TT, MacCormick AD, Arroll B, et al. Text messaging improves preoperative exercise in patients undergoing bariatric surgery. ANZ J Surg 2018 Jun 25:733-738. [doi: 10.1111/ans.14418] [Medline: 29943447]
28. Mundi MS, Lorentz PA, Grothe K, Kellogg TA, Collazo-Clavell ML. Feasibility of smartphone-based education modules and ecological momentary assessment/intervention in pre-bariatric surgery patients. Obes Surg 2015 Oct;25(10):1875-1881. [doi: 10.1007/s11695-015-1617-7] [Medline: 25702141]
29. van der Meij E, Bouwsma EV, van der Heuvel B, Bonjer HJ, Anema JR, Huirne JA. Using e-health in perioperative care: a survey study investigating shortcomings in current perioperative care and possible future solutions. BMC Surg 2017 May 23;17(1):61 [FREE Full text] [doi: 10.1186/s12893-017-0254-6] [Medline: 28535763]
30. Messiah SE, Sacher PM, Yudkin J, Ofori A, Qureshi FG, Schneider B, et al. Application and effectiveness of eHealth strategies for metabolic and bariatric surgery patients: a systematic review. Digital Health 2020;6:2055207619898987. [doi: 10.1177/2055207619898987] [Medline: 32030193]
31. Wolfstadt JI, Soong C, Ward SE. Improving patient outcomes following total joint arthroplasty: is there an app for that? BMJ Qual Saf 2019 Oct 20;28(10):775-777. [doi: 10.1136/bmjqs-2019-009571] [Medline: 31110141]
32. Lee MK, Yun YH, Park H, Lee ES, Jung KH, Noh D. A Web-based self-management exercise and diet intervention for breast cancer survivors: pilot randomized controlled trial. Int J Nurs Stud 2014 Dec;51(12):1557-1567. [doi: 10.1016/j.ijnurstu.2014.04.012] [Medline: 24856854]
33. Grocott MP, Plumb JO, Edwards M, Fecher-Jones I, Levett DZ. Re-designing the pathway to surgery: better care and added value. Perioper Med (Lond) 2017;6:9 [FREE Full text] [doi: 10.1186/s13741-017-0065-4] [Medline: 28649376]
34. Coldebella B, Armfield NR, Bambling M, Hansen J, Edirippulige S. The use of telemedicine for delivering healthcare to bariatric surgery patients: a literature review. J Telemed Telecare 2018 Oct 22;24(10):651-660. [doi: 10.1177/1357633x18795356]
35. Wang CD, Rajaratnam T, Stall B, Hawa R, Sockalingam S. Exploring the Effects of Telemedicine on Bariatric Surgery Follow-up: a Matched Case Control Study. OBES SURG 2019 May 27;29(8):2704-2706. [doi: 10.1007/s11695-019-03930-4]
36. Sherf-Dagan S, Hod K, Mardy-Tilbor L, Gliksman S, Ben-Porat T, Sakran N, et al. The effect of pre-surgery information online lecture on nutrition knowledge and anxiety among bariatric surgery candidates. Obes Surg 2018 Feb 17;28(7):1876-1885. [doi: 10.1007/s11695-018-3134-y]
37. Spanò R, Di Paola N, Bova M, Barbarino A. Value co-creation in healthcare: evidence from innovative therapeutic alternatives for hereditary angioedema. BMC Health Serv Res 2018 Jul 20;18(1):571-513 [FREE Full text] [doi: 10.1186/s12913-018-3389-y] [Medline: 30029666]

38. Janamian T, Crossland L, Wells L. On the road to value co-creation in health care: the role of consumers in defining the destination, planning the journey and sharing the drive. Med J Aust 2016 Apr 18;204(7 Suppl):12-14. [doi: 10.5694/mja16.00123] [Medline: 27078786]
39. Hardyman W, Daunt KL, Kitchener M. Value co-creation through patient engagement in health care: a micro-level approach and research agenda. Public Manag Rev 2014 Feb 07;17(1):90-107. [doi: 10.1080/14719037.2014.881539]
40. Robert G, Cornwell J, Locock L, Purushotham A, Sturmey G, Gager M. Patients and staff as codesigners of healthcare services. Br Med J 2015 Feb 10;350(feb10 14):g7714. [doi: 10.1136/bmj.g7714] [Medline: 25670179]
41. Donetto S, Pierri P, Tsianakas V, Robert G. Experience-based co-design and healthcare improvement: realizing participatory design in the public sector. Design J 2015 May 07;18(2):227-248. [doi: 10.2752/175630615X14212498964312]
42. Witteman HO, Vaisson G, Provencher T, Dansokho SC, Colquhoun H, Dugas M, et al. An 11-item measure of user- and human-centered design for personal health tools (UCD-11): development and validation. J Med Internet Res 2021 Mar 16;23(3):e15032 [FREE Full text] [doi: 10.2196/15032] [Medline: 33724194]
43. Prince RM, Soung Yee A, Parente L, Enright KA, Grunfeld E, Powis M, et al. User-centered design of a web-based tool to support management of chemotherapy-related toxicities in cancer patients. J Med Internet Res 2019 Mar 28;21(3):e9958 [FREE Full text] [doi: 10.2196/jmir.9958] [Medline: 30920373]
44. Korpershoek YJ, Hermsen S, Schoonhoven L, Schuurmans MJ, Trappenburg JC. User-centered design of a mobile health intervention to enhance exacerbation-related self-management in patients with chronic obstructive pulmonary disease (Copilot): mixed methods study. J Med Internet Res 2020 Jun 15;22(6):e15449 [FREE Full text] [doi: 10.2196/15449] [Medline: 32538793]
45. Ledel Solem IK, Varsi C, Eide H, Kristjansdottir OB, Børøsund E, Schreurs KM, et al. A user-centered approach to an evidence-based electronic health pain management intervention for people with chronic pain: design and development of EPIO. J Med Internet Res 2020 Jan 21;22(1):e15889 [FREE Full text] [doi: 10.2196/15889] [Medline: 31961331]
46. Paton C, Kushniruk AW, Borycki EM, English M, Warren J. Improving the usability and safety of digital health systems: the role of predictive human-computer interaction modeling. J Med Internet Res 2021 May 27;23(5):e25281 [FREE Full text] [doi: 10.2196/25281] [Medline: 34042590]
47. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
48. Obesity: clinical assessment and management. National Institute for Health and Care Excellence (NICE). 2016. URL: https://www.nice.org.uk/guidance/qs127/chapter/quality-statement-6-follow-up-care-after-bariatric-surgery: [accessed 2020-06-06]
49. Urquhart C. Grounded theory for qualitative research. In: Grounded Theory for Qualitative Research: A Practical Guide. London: Sage Publications; Nov 06, 2012.
50. Birks M, Mills J. Grounded Theory: A Practical Guide. London: Sage Publications; Feb 09, 2015.
51. Chesnay M. Nursing research using grounded theory: qualitative designs and methods. In: Generating Theory Using Grounded Theory Methodology. New York: Springer; Dec 05, 2014:19-28.
52. Lauti M, Kularatna M, Pillai A, Hill AG, MacCormick AD. A randomised trial of text message support for reducing weight regain following sleeve gastrectomy. Obes Surg 2018 Aug;28(8):2178-2186. [doi: 10.1007/s11695-018-3176-1] [Medline: 29500678]
53. Graham YN, Hayes C, Mahawar KK, Small PK, Attala A, Seymour K, et al. Ascertaining the place of social media and technology for bariatric patient support: what do allied health practitioners think? Obes Surg 2017 Jul;27(7):1691-1696. [doi: 10.1007/s11695-016-2527-z] [Medline: 28054297]
54. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
55. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res in Psychol 2020 Aug 12;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
56. Nelson SJ, Allkins S. Technology in healthcare: the NHS app. Br J Cardiac Nurs 2020 Jan 02;15(1):1-3. [doi: 10.12968/bjca.2019.0147]
57. Diestelkamp S, Wartberg L, Kaess M, Bauer S, Rummel-Kluge C, Becker K, ProHEAD consortium. Effectiveness of a web-based screening and brief intervention with weekly text-message-initiated individualised prompts for reducing risky alcohol use among teenagers: study protocol of a randomised controlled trial within the ProHEAD consortium. Trials 2019 Jan 22;20(1):73 [FREE Full text] [doi: 10.1186/s13063-018-3160-z] [Medline: 30670102]
58. Cook W. The effect of personalised weight feedback on weight loss and health behaviours: evidence from a regression discontinuity design. Health Econ 2019 Jan 27;28(1):161-172. [doi: 10.1002/hec.3829] [Medline: 30260052]
59. Wongvibulsin S, Martin SS, Saria S, Zeger SL, Murphy SA. An individualized, data-driven digital approach for precision behavior change. Am J Lifestyle Med 2020 Apr 25;14(3):289-293 [FREE Full text] [doi: 10.1177/1559827619843489] [Medline: 32477031]
60. Sabharwal A, Fields S, Hilliard M, DeSalvo D. Digital technologies to support behavior change: challenges and opportunities. In: Klonoff DC, Kerr D, Mulvaney SA, editors. Diabetes Digital Health. Amsterdam: Elsevier; Jan 1, 2020:37-50.

61. Lupton D. The digitally engaged patient: self-monitoring and self-care in the digital health era. Soc Theory Health 2013 Jun 19;11(3):256-270. [doi: 10.1057/sth.2013.10]
62. De Vito DA, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
63. King WC, Hsu JY, Belle SH, Courcoulas AP, Eid GM, Flum DR, et al. Pre- to postoperative changes in physical activity: report from the longitudinal assessment of bariatric surgery-2 (LABS-2). Surg Obes Relat Dis 2012;8(5):522-532 [FREE Full text] [doi: 10.1016/j.soard.2011.07.018] [Medline: 21944951]
64. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR Mhealth Uhealth 2016;4(2):e35 [FREE Full text] [doi: 10.2196/mhealth.5120] [Medline: 27076486]
65. Gualtieri L, Rosenbluth S, Phillips J. Can a free wearable activity tracker change behavior? The impact of trackers on adults in a physician-led wellness group. JMIR Res Protoc 2016 Nov 30;5(4):e237 [FREE Full text] [doi: 10.2196/resprot.6534] [Medline: 27903490]
66. Phillips SM, Courneya KS, Welch WA, Gavin KL, Cottrell A, Nielsen A, et al. Breast cancer survivors' preferences for mHealth physical activity interventions: findings from a mixed methods study. J Cancer Surviv 2019 Apr;13(2):292-305. [doi: 10.1007/s11764-019-00751-3] [Medline: 30912011]
67. Nguyen NH, Hadgraft NT, Moore MM, Rosenberg DE, Lynch C, Reeves MM, et al. A qualitative evaluation of breast cancer survivors' acceptance of and preferences for consumer wearable technology activity trackers. Support Care Cancer 2017 Nov;25(11):3375-3384. [doi: 10.1007/s00520-017-3756-y] [Medline: 28540402]
68. Puszkiewicz P, Roberts AL, Smith L, Wardle J, Fisher A. Assessment of cancer survivors' experiences of using a publicly available physical activity mobile application. JMIR Cancer 2016 May 31;2(1):e7 [FREE Full text] [doi: 10.2196/cancer.5380] [Medline: 28410168]
69. Kokts-Porietis RL, Stone CR, Friedenreich CM, Froese A, McDonough M, McNeil J. Breast cancer survivors' perspectives on a home-based physical activity intervention utilizing wearable technology. Support Care Cancer 2019 Aug;27(8):2885-2892. [doi: 10.1007/s00520-018-4581-7] [Medline: 30554276]
70. Hardcastle SJ, Galliott M, Lynch BM, Nguyen NH, Cohen PA, Mohan GR, et al. Acceptability and utility of, and preference for wearable activity trackers amongst non-metropolitan cancer survivors. PLoS ONE 2018 Dec 31;13(12):e0210039. [doi: 10.1371/journal.pone.0210039]
71. Kalarchian MA, Marcus MD, Courcoulas AP, Cheng Y, Levine MD. Preoperative lifestyle intervention in bariatric surgery: initial results from a randomized, controlled trial. Obesity (Silver Spring) 2013 Feb;21(2):254-260 [FREE Full text] [doi: 10.1002/oby.20069] [Medline: 23404832]
72. Robinson A, Slight R, Husband A, Slight S. The value of teachable moments in surgical patient care and the supportive role of digital technologies. Perioper Med 2020;9:1-3 [FREE Full text] [doi: 10.1186/s13741-019-0133-z] [Medline: 32042404]
73. Flocke SA, Clark E, Antognoli E, Mason MJ, Lawson PJ, Smith S, et al. Teachable moments for health behavior change and intermediate patient outcomes. Patient Educ Couns 2014 Jul;96(1):43-49 [FREE Full text] [doi: 10.1016/j.pec.2014.03.014] [Medline: 24856449]
74. Cable TA, Meland E, Søberg T, Slagsvold S. Lessons from the Oslo study diet and anti-smoking trial: a qualitative study of long-term behaviour change. Scand J Public Health 2016 Sep 05;27(3):206-212. [doi: 10.1177/14034948990270031301]
75. Lepping P, Palmstierna T, Raveesh BN. Paternalism v. autonomy - are we barking up the wrong tree? Br J Psychiatry 2016 Aug 02;209(2):95-96. [doi: 10.1192/bjp.bp.116.181032] [Medline: 27482035]
76. Curry SJ, Wagner EH, Grothaus LC. Evaluation of intrinsic and extrinsic motivation interventions with a self-help smoking cessation program. J Consult Clin Psychol 1991 Apr;59(2):318-324. [doi: 10.1037//0022-006x.59.2.318] [Medline: 2030194]
77. Patrick H, Canevello A. Methodological overview of a self-determination theory-based computerized intervention to promote leisure-time physical activity. Psychol Sport Exerc 2011 Jan;12(1):13-19 [FREE Full text] [doi: 10.1016/j.psychsport.2010.04.011] [Medline: 21103069]
78. Williams GG, Gagné M, Ryan RM, Deci EL. Facilitating autonomous motivation for smoking cessation. Health Psychol 2002;21(1):40-50. [doi: 10.1037/0278-6133.21.1.40]
79. Williams GC, Grow VM, Freedman ZR, Ryan RM, Deci EL. Motivational predictors of weight loss and weight-loss maintenance. J Personal Soc Psychol 1996;70(1):115-126. [doi: 10.1037/0022-3514.70.1.115]
80. Yang C, Hui Z, Zeng D, Liu L, Lee DT. Examining and adapting the information-motivation-behavioural skills model of medication adherence among community-dwelling older patients with multimorbidity: protocol for a cross-sectional study. BMJ Open 2020 Mar 24;10(3):e033431 [FREE Full text] [doi: 10.1136/bmjopen-2019-033431] [Medline: 32209623]
81. Shumlich EJ, Fisher WA. An information-motivation-behavioural skills model analysis of young adults' sexual behaviour patterns and regulatory requirements for sexual consent in Canada. Can J Hum Sexual 2019 Nov 01;28(3):277-291. [doi: 10.3138/cjhs.2018-0040]
82. Rongkavilit C, Naar-King S, Kaljee LM, Panthong A, Koken JA, Bunupuradah T, et al. Applying the information-motivation-behavioral skills model in medication adherence among Thai youth living with HIV: a qualitative

study. AIDS Patient Care STDS 2010 Dec;24(12):787-794 [FREE Full text] [doi: 10.1089/apc.2010.0069] [Medline: 21091238]
83. Chang S, Choi S, Kim S, Song M. Intervention strategies based on information-motivation-behavioral skills model for health behavior change: a systematic review. Asian Nurs Res 2014 Sep;8(3):172-181 [FREE Full text] [doi: 10.1016/j.anr.2014.08.002]
84. Gulmans J, Vollenbroek-Hutten M, van Gemert-Pijnen L, van Harten W. A web-based communication system for integrated care in cerebral palsy: experienced contribution to parent-professional communication. Int J Integr Care 2012;12:e9 [FREE Full text] [doi: 10.5334/ijic.672] [Medline: 22977435]
85. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
86. Vooijs M, Alpay LL, Snoeck-Stroband JB, Beerthuizen T, Siemonsma PC, Abbink JJ, et al. Validity and usability of low-cost accelerometers for internet-based self-monitoring of physical activity in patients with chronic obstructive pulmonary disease. Interact J Med Res 2014 Oct 27;3(4):e3056 [FREE Full text] [doi: 10.2196/ijmr.3056] [Medline: 25347989]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
NHS: National Health Service
SDT: Self-determination Theory

Edited by A Kushniruk; submitted 20.04.21; peer-reviewed by C Richardson, S Goldstein; comments to author 26.08.21; revised version received 05.10.21; accepted 06.11.21; published 04.03.22.

Please cite as:
Robinson A, Husband A, Slight R, Slight SP
Designing Digital Health Technology to Support Patients Before and After Bariatric Surgery: Qualitative Study Exploring Patient Desires, Suggestions, and Reflections to Support Lifestyle Behavior Change
JMIR Hum Factors 2022;9(1):e29782
URL: https://humanfactors.jmir.org/2022/1/e29782
doi:10.2196/29782
PMID:35254271

Original Paper

Mobile Health to Support Community-Integration of Individuals With Disabilities Using iMHere 2.0: Focus Group Study

Rebecca E Ward[1], MD; I Made Agus Setiawan[2,3], MSc; Eleanor Quinby[1,4], BS; Melva Fair[5], MS; Zara Ambadar[1,2], PhD; Bambang Parmanto[2], PhD; Brad E Dicianno[1,4], MD

[1]Department of Physical Medicine and Rehabilitation, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[2]Department of Health Information Management, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA, United States

[3]Department of Computer Science, Udayana University, Badung, Bali, Indonesia

[4]Human Engineering Research Laboratories, Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, PA, United States

[5]Community Living and Support Services, Pittsburgh, PA, United States

Corresponding Author:
Brad E Dicianno, MD
Human Engineering Research Laboratories
Veterans Affairs Pittsburgh Healthcare System
VA Pittsburgh Healthcare System
6425 Penn AVE, Suite 400
Pittsburgh, PA, 15206-4022
United States
Phone: 1 412 822 3700
Email: dicianno@pitt.edu

Abstract

Background: Mobile health (mHealth) systems that support self-management can improve medical, functional, and psychosocial outcomes for individuals with disabilities and chronic conditions. The mHealth systems can potentially be expanded to support community integration.

Objective: The purposes of this study were to (1) partner with a community-based organization that supports community integration of individuals with disabilities; (2) identify software requirements needed to support community participation; and (3) iteratively refine an existing mHealth application to include new requirements.

Methods: Community Living and Support Services (CLASS), a nonprofit organization that serves individuals with disabilities in Pittsburgh, Pennsylvania, was identified as the focus group for this study. Key stakeholders within the Community Partners Program at CLASS proposed design requirements for an existing mHealth application, Interactive Mobile Health and Rehabilitation (iMHere) 2.0, that has been used to support self-management.

Results: We gathered qualitative data from a focus group composed of CLASS members to develop and iteratively revise iMHere 2.0 to include new modules and features to support community integration. A caregiver app was also developed. The new system contains features to support finance, transportation, client and caregiver communication, calendar and checklist management, upcoming medical and nonmedical appointments, social engagement, pain management, and access to a personal profile. Modifications were made to the following existing modules: education, mood, personal health record, goals, medications, and nutrition.

Conclusions: A successful partnership with a community-based organization that supports individuals with disabilities resulted in a newly designed mHealth system with features to support community integration.

(JMIR Hum Factors 2022;9(1):e31376) doi:10.2196/31376

KEYWORDS

community integration; self-care; mobile health; smartphone; rehabilitation; disability; mobile phone

Introduction

Living in the community affords many benefits for individuals with disabilities and chronic conditions. Studies have shown that higher integration into the community is associated with better health outcomes, longevity, higher quality of life, and lower cost of care [1-3]. However, many barriers to community living exist, many of which are due to fragmented outpatient, inpatient, long-term, and community-based services operating under different funding streams and regulations. This system isolates individuals, leaves many medical and social needs unmet, and ultimately drives costs to enormous and unsustainable levels [4]. It has been proposed that software-based tools—specifically mobile health (mHealth) systems—may be able to improve the process [5].

Background

Interactive Mobile Health and Rehabilitation (iMHere) is an mHealth system designed to help individuals, including those with disabilities, manage chronic and complex medical conditions so that they can live independently and integrate more fully into the community. The original version of iMHere (iMHere 1.0) focused primarily on self-management tasks aimed to prevent medical complications and promote health. Five modules were developed to allow the user to manage skin care, medications, bladder self-catheterization, bowel programs, and mental health [6]. Through these modules, users could indicate completion of self-management tasks, send adherence data to a portal monitored by a clinician, and receive personalized regimens from a clinician. By providing a secure connection between a client's smartphone and a web-based clinician portal, iMHere supported Health Insurance Portability and Accountability Act–compliant messaging between a client and team of clinicians [7]. The portal categorized and organized client responses in a dashboard that triaged and flagged data based on urgency.

Prior Work

Preliminary research with iMHere 1.0 includes a series of focus groups, usability studies, accessibility studies, and clinical trials. Focus groups and usability studies have driven an iterative, user-centered design process and demonstrated that individuals desire an app that is easy to use and engaging, that can provide educational materials, motivation, and support, and that can be personalized [8-13]. Accessibility studies resulted in improvements to accessibility for persons with motor, sensory, and cognitive impairments [14-16]. A clinical trial demonstrated that individuals with spina bifida who were frequent users of iMHere showed positive changes in self-management as well as reduced need for caregiver support [17]. iMHere also demonstrated potential for monitoring and preventing skin wounds in individuals with spina bifida [18]. A recent study found that the iMHere mHealth system aided in the prevention of urinary tract infections and reduction in depression symptoms in individuals with spinal cord injury [19].

Based on this prior research, iMHere 2.0 was developed [11]. The new features include dual iOS and Android platforms, new accessibility and personalization features, a reward center, and 7 additional modules (education, personal health record [PHR], goal setting, wheelchair maintenance, exercise, nutrition, and supplies). Further, a caregiver app allows family members, legal guardians, and attendant care providers to monitor the client's activity and provide encouragement.

Study Goal

The primary aim of this project was to expand iMHere 2.0's functionality to support a community-based organization that assists individuals with disabilities and encourages independent living. To live independently in the community, individuals with disabilities and chronic conditions often need services and supports provided by community-based organizations. We partnered with a community-based organization that fosters community inclusion for individuals with disabilities. Located in the East End of Pittsburgh, Pennsylvania, Community Living and Support Services (CLASS) is a nonprofit organization that offers a variety of individualized services such as community-based case management for social, recreational, and residential supports for individuals with disabilities. This manuscript describes the iterative development process of the features created for iMHere to support services offered by CLASS through the examination of qualitative data obtained from a focus group.

Methods

The study was approved by the University of Pittsburgh Institutional Review Board (STUDY20020049), and all participants underwent an informed consent process. Key stakeholders who oversee the Community Partners Program (CPP) at CLASS, an initiative that assists individuals with disabilities in both their homes and communities, served as this study's focus group.

Individuals with disabilities enrolled in the CPP choose specific goals and develop strategies to meet their needs. This program provides one-on-one support. Each client is assigned a case manager who supports short- or long-term goals based on the individual's defined needs. The CPP emphasizes community integration and strives to reduce the amount of both medical and nonmedical services required by clients. Further, the CPP can partner with individuals to foster connections with community resources, assist with decision-making and problem-solving, impart compensatory strategies, help in the search for employment and volunteer opportunities, and provide life skills training.

With the aims of the CPP in mind, we met with the CPP stakeholders on January 26, 2016, and July 6, 2017, to discuss how iMHere 2.0 could best support a community-based organization and to subsequently develop design criteria. At these initial focus group meetings, stakeholders were given a demonstration of how iMHere 2.0 functions. Stakeholders then presented the documents and forms used by case managers and other employees to manage a client's case. Information from these documents was organized into themes and used to develop a mock-up for a new design of iMHere. This mock-up was presented to the stakeholders on March 26, 2018. Feedback was gathered and subsequent changes to iMHere 2.0 were made and presented again on April 5, 2019.

Results

Overview

A focus group of 15 stakeholders participated in meetings. The demographics of the stakeholders are presented in Table 1 to illustrate the diversity of the participants and the roles that they would play in using the iMHere 2.0 system. The participants represented a variety of roles at CLASS and had different levels of interaction with clients. The focus group was mainly composed of participants who had a professional relationship with the client. Table 2 outlines the demographics of the clients who received services from the focus group participants. Some stakeholders were assigned to more than one client and, therefore, played more than one role.

Stakeholders identified 5 roles for iMHere 2.0 users based on existing CPP workflow. The "client" is defined as the person receiving services through CLASS. These users would have access to the client app. They can receive services from one or more of the care roles. "Case managers" include program directors and those who oversee the 3 other care roles. Case managers would be web portal users with the highest level of access to the system. "Service workers" include community partners who work with the client to develop and meet goals regarding self-care. Service workers oversee attendants and directly track client activity. These users would have access to both the web portal and the caregiver app. "Caregivers" are those who provide assistance to the client in an unpaid role and would have access to the caregiver app. This category can include family members or friends who provide direct care to clients and need to directly track client activities or ensure that clients complete daily tasks. The "attendant" role is defined as a paid, direct care worker who needs to keep a record of the assigned client's completed tasks. Attendants would have access only to the caregiver app, with an attendant-specific module. Attendants can serve more than one client and are expected to use the app primarily for recording task completion, which can be approved by the client. Figure 1 shows how users within each role would utilize the 3 frontend components of iMHere 2.0: Web Portal, Client App, and Caregiver App.

Stakeholders requested many new features for iMHere 2.0 (Multimedia Appendix 1). The qualitative data collected from the focus group were bundled into common themes. Features with higher priority (necessary) were developed for the current system, while those with lower priority (desired) were set as future design criteria. Some new features were implemented as additions to existing modules or the portal, while others were built as new modules or portal features.

Table 1. Demographics of focus group participants (N=15).

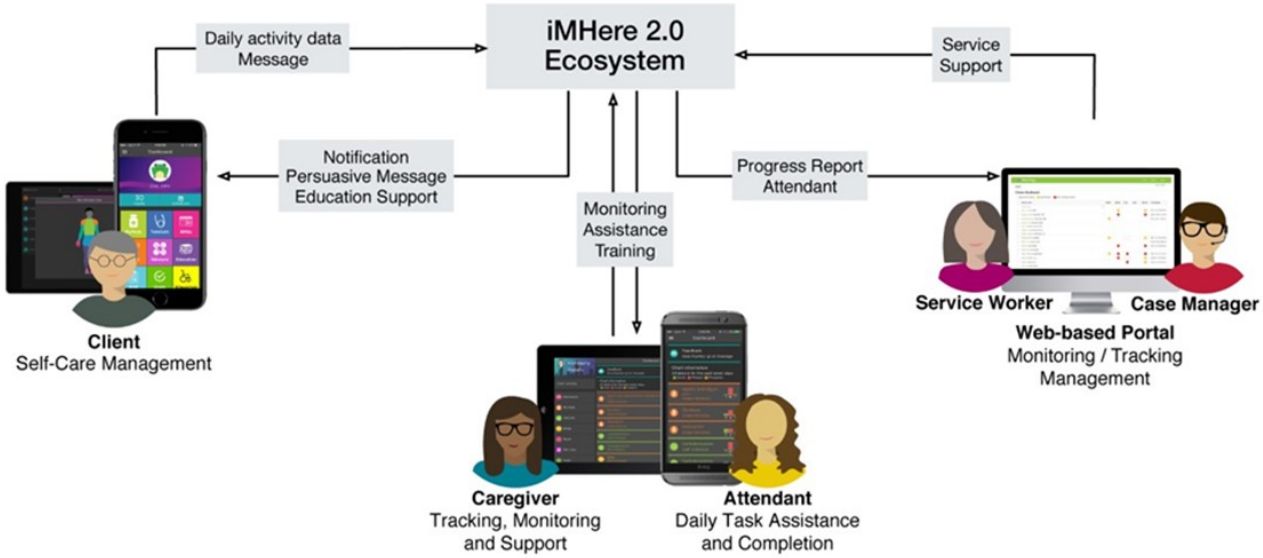
Characteristic	n
Age (years)	
25-34	5
35-44	1
45-54	7
55-64	1
65-74	1
Gender	
Female	14
Male	1
Race	
White/Caucasian	14
Black/African American	1
Ethnicity	
Hispanic/Latinx	13
Non-Hispanic/not Latinx	2
Household income	
25,000 or less	1
26,000-50,000	6
51,000-75,000	3
76,000-100,000	1
101,000-125,000	1
More than 125,000	1
Declined to answer	2
Community of residence	
Suburban	13
Urban	1
Rural	1
Education level	
High school diploma	2
Trade school diploma	1
Associates degree	5
Bachelor’s degree	6
Master’s degree	1
Marital status	
Married/long-term relationship	10
Expected role in the iMHere[a] system (some participants had more than one role)	
Portal user-administrator (technical support or training)	3
Portal user-service worker	7
Portal user-speech language pathologist	1
Portal user-case manager	5
CG app user-direct care provider/attendant care worker	4

[a]iMHere: Interactive Mobile Health and Rehabilitation.

Table 2. Demographics of clients served by stakeholders.

Characteristic	n
Clients served	
Male clients	6
Female clients	6
Both male and female clients	3
Client age (years)	
18-24	2
25-34	2
35-44	2
45-54	4
55-64	8
65-74	0
≥75	1
Client diagnosis	
Cerebral palsy	8
Traumatic brain injury	5
Spina bifida	2
Spinal cord injury	1
Autism spectrum disorder/intellectual disability	1
Diabetes and heart disease	1
Client secondary diagnosis	
Medical	3
Mental health	2
Learning disability	1
Multiple	3
None	6

Figure 1. Overview of the iMHere 2.0 system.



Caregiver App

The new caregiver app is designed to support a variety of common relationships between the caregiver and client (Figure 2). In previous iterations of the app, the functionality targeted only the unpaid (family) caregiver model. As shown in Figure 2A, the updated design of this app is now customizable to allow for meaningful use by both unpaid caregivers and paid caregivers. The new caregiver app includes tools that simplify the unpaid caregiver's monitoring activities while also allowing access to client reports through each module (Figure 2B). For the paid caregiver, or attendant, the app retains monitoring and feedback functionality, permitting attendants to virtually track their clients' activity as needed (Figure 2C-E).

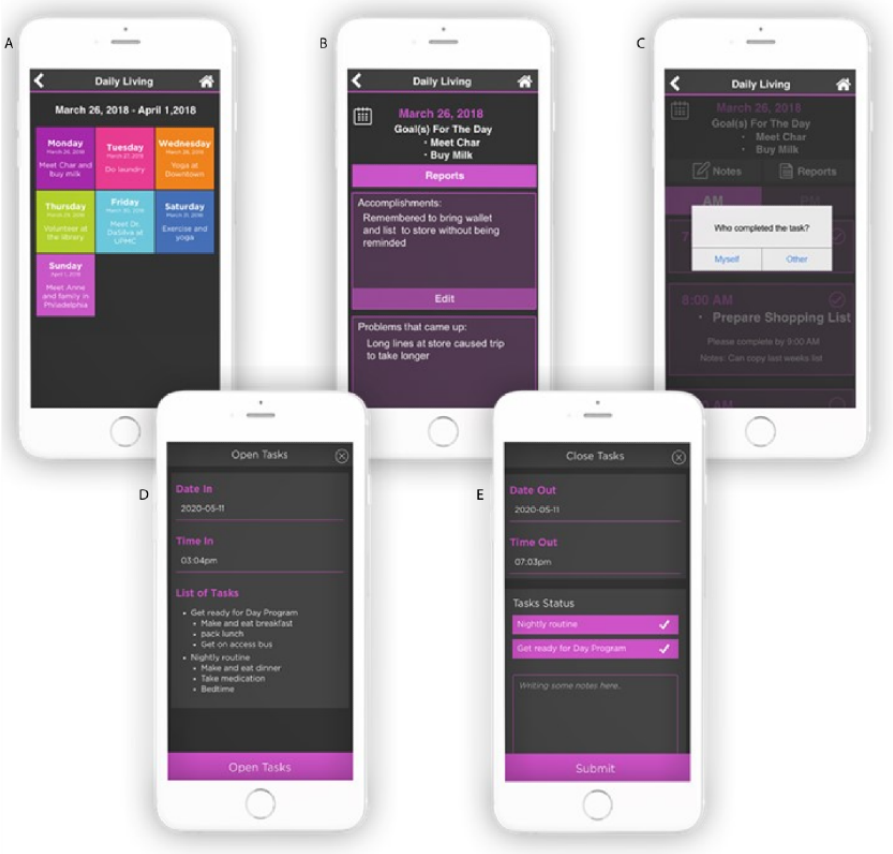
Figure 2. The caregiver app that allows attendants and other caregivers to monitor and provide feedback to their client/family member and access summary and timeline views.

Monitoring or Tracking

CLASS suggested several monitoring or tracking changes to the app, which resulted in the creation of the Attendant Log module. Through free-text spaces and progress note areas in the Attendant Log module, attendants and caregivers track the progress of care provided to the client. Within the Attendant Log module, CLASS can electronically verify visits and track attendant activities such as "clocking-in" and "clocking-out" of sessions and completion of daily tasks.

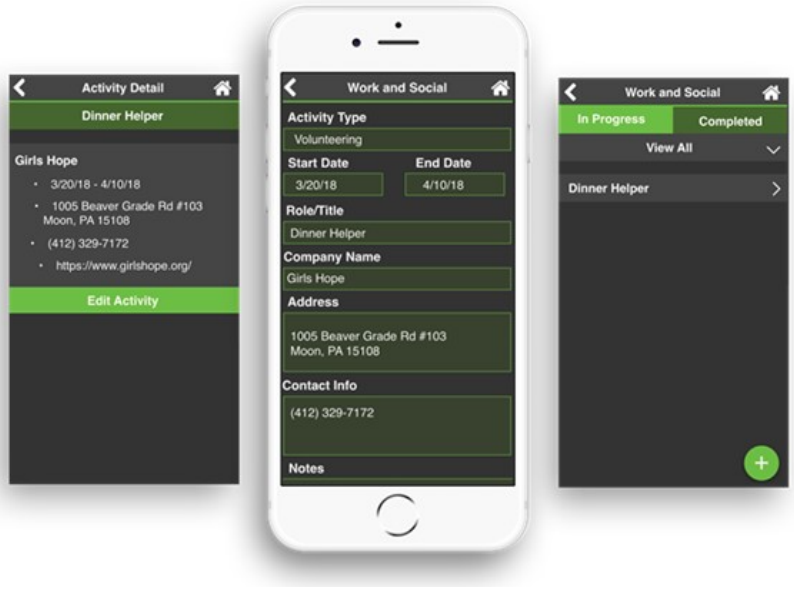
The Daily Living module, another addition created because of the focus group's suggestions, provides both a weekly schedule and checklist for clients. Figure 3A shows the landing screen of the module, from which clients can view and edit their errands, training sessions, appointments, and other events for each day of the week. Clients can create goals for the day, enter notes, and view reports of their activity. As shown in Figure 3B, Goal reports list the client's accomplishments and any encountered problems. Figure 3C shows that on completion of a task (eg, buying milk), the client is prompted to enter "who completed the task," with options of selecting "myself" or "other." "Other" typically refers to the client's attendant because clients can specify difficult tasks for the caregiver to complete. These tasks appear on a checklist in the caregiver app. The attendant can record the day's activities in the caregiver app, while the client can confirm and add notes. Caregivers can track tasks, and attendants can complete tasks within a module created in the caregiver app. Attendants can "clock in" and complete tasks as shown in Figure 3D-E. When attendants "clock out," the client can confirm that tasks were indeed completed. This feature was developed because the focus group identified the need for improved attendant accountability and communication with clients. The goal of the client confirmation feature is to provide a second layer of authentication to information provided by the attendant about completion of tasks and time of arrival and departure.

Figure 3. The Daily Living module that allows clients to see their upcoming events for the week. Daily goal reports display the client's accomplishments and any encountered problems. Attendants can clock in and clock out, access task lists, and document completion of assigned tasks.



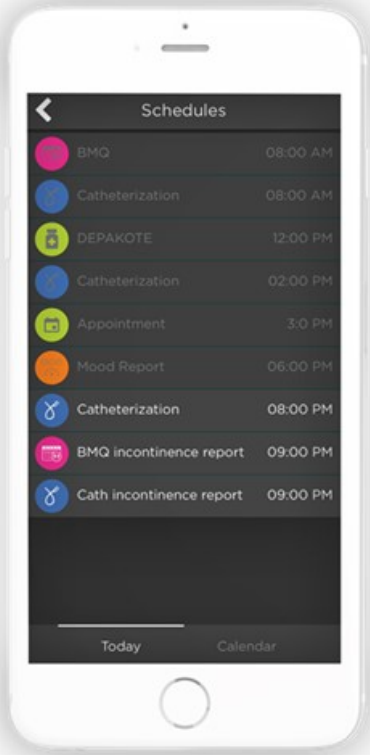
Within the newly created Work and Social module, clients can enter volunteer, work, and social activities along with start and end dates, roles or titles, company names, addresses, contact information, and notes. Clients mark their activities as "in progress" or "completed" (Figure 4).

Figure 4. The Work and Social module that allows clients to enter details about work, social, and volunteer activities.



CLASS recommended several monitoring and tracking changes to pre-existing modules. A medication adherence tracker was added to the My Meds module. The tracker allows caregivers and attendants to track client medication adherence. Now, through a centralized calendar (Figure 5), clients have access to reminders and medical and nonmedical appointments. Under the pre-existing Exercise module, clients can now enter their workouts and track their exercise progress and activity.

Figure 5. The calendar and schedule features that allow clients to track activities, appointments, and events.



Through a future Incident Tracking module, clients will be able to report falls and related injuries. A Pain Tracker will also be included so that clients can document pain characteristics. It will also be possible to track attendant training from the portal. An electronic record documenting that the caregiver reviewed training information and complied with funding sources will automatically be created and stored in both the clinician portal and a private cloud. The system will also track what modules the attendant completed. Finally, clients will be able to track expenses and learn how to budget in a future finance module.

Education

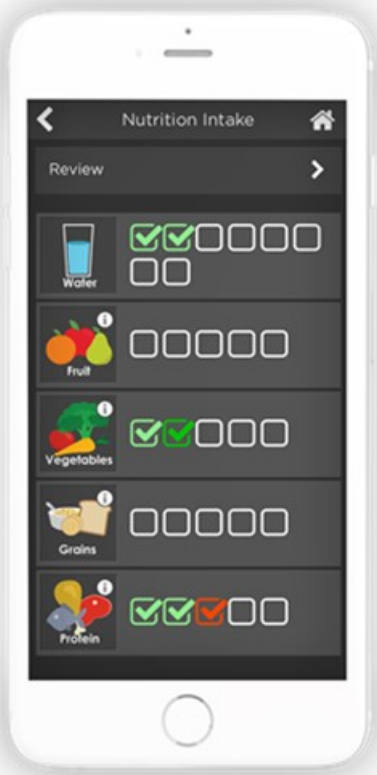
CLASS requested the inclusion of a new Training module. The newly designed training module contains work-related and non–work-related training. Both active and past training are recorded. Within the module, clients can enter the job training location, addresses, and contact information (Figure 6). Clients can note their work progress and develop follow-up plans.

Figure 6. The Training module that allows clients to enter work-related and non–work-related training information.

CLASS also suggested updating the pre-existing Nutrition module (Figure 7) to provide more detailed diet information, including specific recommendations, such as recipes, to support healthy eating. In response to their comments, the future nutrition module will include recipe recommendations, a list of health foods, and information on various diets (eg, low-cholesterol diets, diets for patients with diabetes, gluten-free diets). Further, CLASS requested that certain education modules be available for caregivers and attendants. Through these modules, caregivers and attendants will be able to learn more about their clients' conditions, train on the client's care plan, and track client results of the quiz for content retention. These changes will be implemented in future versions.

The CLASS focus group also requested that a module be designed to provide wheelchair users with more information on wheelchair features, maintenance, and training. Within the wheelchair module, the guidebook was developed for clients to access educational information about power and manual wheelchairs. Under the manual wheelchair section, users can view a list of manual wheelchair components with corresponding images; for example, users can learn about frame types, cushion options, positioning accessories, and handrim designs. Users can also access information on how to ergonomically optimize their wheelchair through the set up and positioning of different components. The manual wheelchair skills section allows users to view videos of various wheelchair skills ranging from simple maneuvers to more advanced skills. Currently, the power wheelchair section allows users to view various component options for power wheelchairs through images and Graphics Interchange Format that illustrate the functionality of each component.

Figure 7. The Nutrition module provides nutritional advice and education.

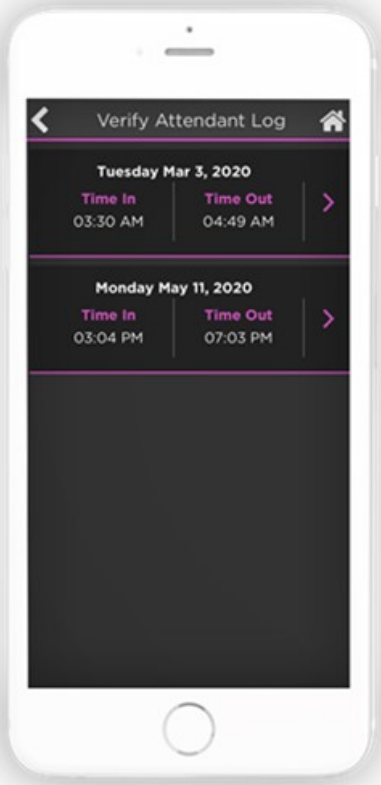


Support

A module that allows clients to input feedback, such as the late arrival of an attendant, was also suggested. The Notes and Reports section within the new Daily Living module allows feedback input; after attendants complete their session, clients can accept or reject their report of completed tasks and note any issues (Figure 8).

CLASS requested changes to the pre-existing Goal Setting module for meaningful reporting of goal progress to funding organizations. The revised Goal Setting module provides users with a centralized area for goal creation. Here, clients are encouraged to include nutrition goals, such as healthy eating and physician-approved diets, as well as goals to increase their independence and community participation. Through the caregiver app, caregivers and service workers can ensure that clients are achieving their specified goals and view tracking data that quantifies goal progress. With respect to a future design criterion, clients will also be able to upload exercise photos to receive feedback and suggestions related to exercise.

Figure 8. Clients can verify that an attendant provided care and completed tasks.



Privacy and Security

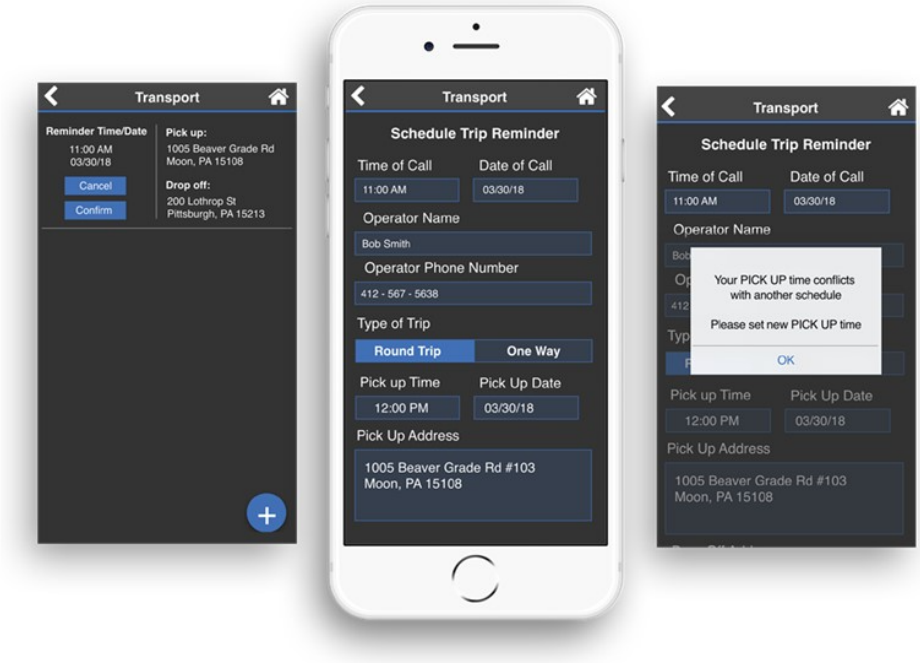
CLASS requested termination settings, such as the ability to remove the app from the client's phone via the portal, which are now included under the Configuration module. Local information stored on the app can now be removed remotely via the portal. Additionally, caregivers can now access information about multiple clients through the Client Switcher module. This is useful for attendants who are usually assigned to multiple clients.

In the future, clients will be able to change their privacy settings to control what personal information is shared with the caregiver. Of note, the client app does not need to be paired with a caregiver app if the client does not have an assigned attendant. With respect to client app parameters, certain features will be locked. For example, it is possible to lock content such that a client is unable to delete medication or exercise reminders if that feature is required for a particular program.

Reminders

Through the newly created Transportation module, clients now receive reminders to arrange for transportation after scheduling an appointment. Within the Transportation module, clients can enter a preferred time for a pick-up reminder call, date of the reminder call, operator name, operator phone number, type of trip (round trip or one way), pick-up time, pick-up date, pick-up address, and drop-off address. Clients receive a warning message if their specified pick-up time conflicts with another calendar activity (Figure 9). If transportation is added by a caregiver, clients must confirm the newly scheduled event. Clients can also cancel their transportation but must enter a reason for cancellation.

Figure 9. The Transportation module that allows clients to enter information pertaining to upcoming transportation.



In the future, through a Global Reminder feature, clients will be periodically reminded of important agenda items, including ad-hoc prompts to report possible needs (eg, "Do you need to go to the grocery store this week?"). The Exercise module will push notifications that encourage exercise to the client. Finally, future design criteria will include "charge device" audio reminders as clients often forget to charge their devices.

CLASS noted that clients would benefit from a Supplies module that will allow users to track their medical supplies and set reminders for maintaining their stock. This module allows clients to enter information about their medical supplies such as type, purpose, vendor, quantity, and order date. In the Supplies module, clients can set reminders to order more supplies as needed. The app includes auto-fill supply options in addition to free-text functions. Future design criteria will allow for more automation, such as monthly reminders that alert users of the need to reorder a 30-day stock of supplies.

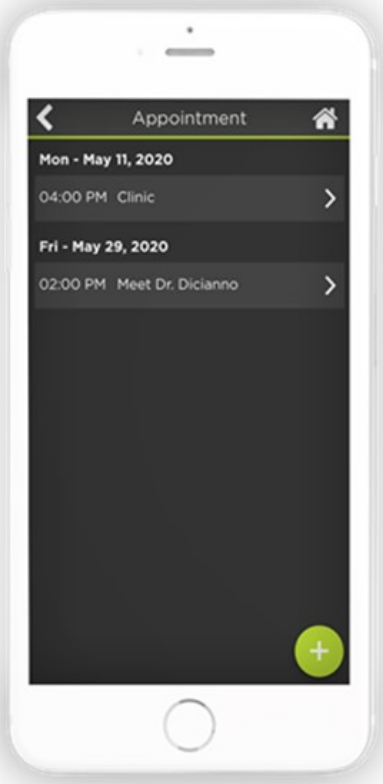
Accessibility

CLASS requested text-to-speech and voice-command features. These requests were not implemented because Apple and Android devices contain built-in accessibility features. These features include "spoken speech/content" that allows users with limited vision to hear text, "voice commands" that allows users with impaired mobility to control their devices with verbal commands, and "switch access" that allows individuals with impaired mobility to control their device using a keyboard or a mouse. Further, iMHere 2.0 already contains accessibility features that allow clients to adjust text size, line height, button size, font style, button spacing, and hand preference. Mood scales consisting of emoticons instead of numbers will be implemented in a future build.

Notes

The Appointment module is a newly created submodule within the PHR module. The Appointment module includes a calendar that displays upcoming medical and nonmedical appointments (Figure 10). By selecting the appointment, the client can view details such as date, time, and provider information (name, address, and phone). Additional fields for a summary of the appointment and notes are also included. The client can enter the appointment details by selecting "New Appointment." Before submitting the form, the client must indicate whether they need to schedule transportation. If clients do not have enough information to complete the form in its entirety, they will receive reminders until the form is completed and submitted. Within the notes section, clients can create a checklist of medical condition–related questions to ask the physician during doctor appointments.

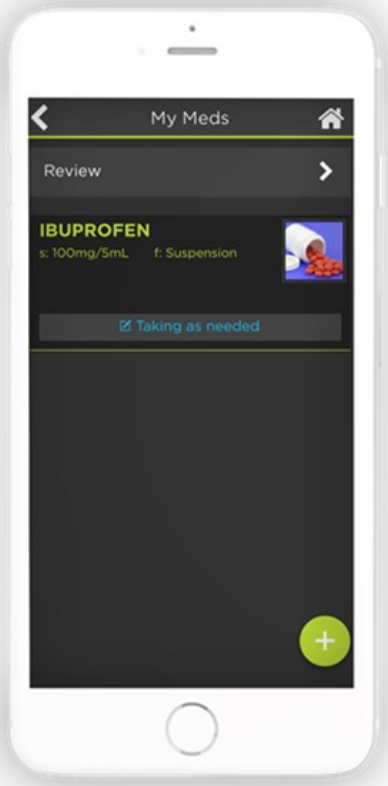
Figure 10. The calendar displays upcoming medical and nonmedical appointments.



CLASS also recommended changes to the pre-existing medication module. To ensure that clients are taking the correct medications, CLASS asked that clients be able to upload pictures of pills and medication containers to the medication module (Figure 11), which was a feature that was added. Both the client and the caregiver also have access to a Medication Tracker that records client medication adherence.

With respect to future direction, clients will ultimately have access to a medical log feature within the pre-existing PHR module. The medical log will include vital signs (blood pressure, heart rate, body temperature, or respiratory rate), blood sugar readings, and height, weight, and body mass index measurements. To enter data, clients, caregivers, or clinicians will be able to add the date of the reading, the time, and the measurements (eg, systolic and diastolic pressures). Clients will eventually be able to view a graphical representation of their readings for a single day, week, month, or entire year.

Figure 11. Pictures of pills and medication containers that allow clients to identify medications.



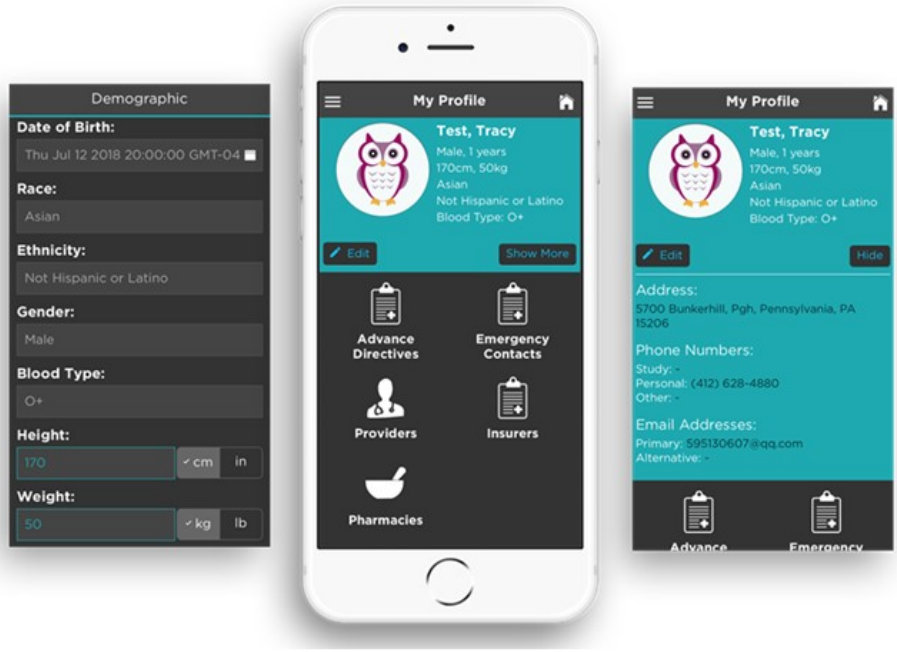
Safety

An emergency contact can be added from the phone's lock screen. In Case of Emergency (ICE) contacts are accessible in the client profile. By selecting ICE, clients can retrieve contact information for their provider, support staff, and insurance company. The ICE module also lists restrictions such as medications or procedures that should be avoided during a medical emergency. Under "My Profile," clients can also find advance directives, additional emergency contacts, and a list of providers and medications. QR code availability in future will allow users to scan their medications to confirm that the correct medication is administered.

Profile

CLASS suggested changes within the PHR module. The focus group proposed that staff could better serve clients if identifying client information was added to the PHR module. Within the revised module, clients can create and view a customized user profile. Clients can enter their gender, age, height, weight, race, ethnicity, and blood type (Figure 12). Clients can also update their profile to include information about likes, dislikes, and other introductory information such as pets at home, smoking status, and medical conditions, which can be then pushed to the attendant. If a client smokes or has a pet at home, the attendant will receive a digital alert.

Figure 12. Clients can enter their contact information, demographics, emergency contacts, advanced directives, insurers, providers, and pharmacies.



Overall Development Themes

Stakeholders emphasized that the revised iMHere 2.0 should remain client-driven to encourage self-sufficiency. They also stated that the system should be useable by the client alone or with a variety of different support roles from caregivers, attendants, and care managers. The original iMHere system 2.0 was designed as an adaptive system that allows clients to receive modules and content only as needed. The stakeholders felt it should continue to provide personalization of modules such that different clients could have different profiles and app configurations depending on their unique needs.

Discussion

Principal Findings

This study demonstrates how stakeholders within a community-based organization that provides services to individuals with disabilities can influence and shape the design of mHealth technology and how these designs can be improved to better serve individuals with disabilities.

The literature on chronic care management shows that promoting partnerships between recipients of services and the providers of the services has positive outcomes, especially when such partnerships facilitate self-management [20,21]. Important self-management skills include problem solving, decision-making, resource utilization, and goal-setting [21]. Thus, innovative systems to support individuals with chronic conditions and disabilities should not only incorporate features that empower users with tools for self-management but also facilitate sustainable partnerships.

mHealth systems are a potential way to deliver such support and empowerment. Some community-based organizations have expressed the need for dashboards that show live data to facilitate care decisions in real time [22] and allow users to record data “on the go” as a way to improve quality metric reporting [23]. An ideal software system could show the interrelationships among an individual’s medical problems, daily schedule, medical and social care plans, and goals [22]. Solutions are also needed to allow care providers to recognize red flag symptoms and intervene quickly, manage medications, and coordinate care while integrating the support of family and other caregivers of individuals with chronic conditions and disabilities [24].

To our knowledge, iMHere 2.0 is the only mHealth system designed to meet these needs and designed iteratively in partnership with the organization it serves. By gathering qualitative data from a focus group, this study was able to identify ways to re-design mHealth technology to address the needs of an organization supporting individuals with disabilities living in the community. This partnership resulted in a highly fruitful development process that led to major advances in the mHealth system. Additionally, our team generated design criteria to be utilized in future iterations of the system.

Some limitations of our development process, however, deserve discussion. First, iMHere 2.0 was not designed to meet the needs of all community-based organizations or users. Concurrent work is being carried out to further refine iMHere 2.0 to support other organizations with different workflows, users, and organizational structure. Second, although we attempted to incorporate all the necessary changes suggested by CLASS, some suggestions could be met with native features of phones or tablets. For instance, low–battery-level reminders, text-to-speech or read-out-loud features, and voice-command features exist. Because these may not meet the needs of all users, future accessibility advancements are needed for users with specific accessibility needs.

Future Work

Ongoing work is being conducted to assess the usability and feasibility of iMHere 2.0 and to study the impact of augmenting

programs offered at CLASS with iMHere 2.0 to measure the impact on client outcomes such as community integration. We are also conducting implementation research with other community-based organizations, including those that provide attendant care services in the home, to identify facilitators and barriers to wide-scale implementation of the technology. In parallel, we will continue development that is inspired by design criteria generated from each of our studies. We also aim to integrate some features of iMHere into the electronic health record system to reduce data entry requirements and facilitate data sharing.

Conclusions

Individuals with disabilities can thrive in community settings when they are given access to high-quality health care, supportive caregivers, and community support. Using an iterative design process in partnership with a community-based organization, we built an mHealth system with new features to support community integration of individuals with disabilities.

Acknowledgments

The contents of this publication were developed under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90DPGE0002). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this publication do not necessarily represent the policy of NIDILRR, ACL, HHS, and endorsement by the Federal Government should not be assumed. The authors thank the many individuals who participated in this research, without whom it would not be possible. They would also like to thank Charlene Raynovich and Dr Al Condeluci for assistance and guidance in carrying out the project and Lauren Byrnes for graphic design.

Conflicts of Interest

BED, BP, and IMAS are inventors of the iMHere mHealth system.

Multimedia Appendix 1
The requested new features for iMHere 2.0.
[DOCX File , 24 KB - humanfactors_v9i1e31376_app1.docx]

References

1. Berkman L, Kawachi I, Glymour M. Social Epidemiology. Oxford, UK: Oxford University Press; Aug 06, 2014:1-640.
2. Putnam RD. Tuning in, tuning out: the strange disappearance of social capital in America. Political Sci Politics 1995 Dec;28(4):664. [doi: 10.2307/420517]
3. Putnam RD. Bowling alone: America's declining social capital. In: The City Reader. London: Routledge; 2015.
4. Voorhees CC, Murray D, Welk G, Birnbaum A, Ribisl KM, Johnson CC, et al. The role of peer social network factors and physical activity in adolescent girls. Am J Health Behav 2005;29(2):183-190 [FREE Full text] [doi: 10.5993/ajhb.29.2.9] [Medline: 15698985]
5. Enguidanos SM, Gibbs NE, Simmons WJ, Savoni KJ, Jamison PM, Hackstaff L, et al. Kaiser Permanente community partners project: improving geriatric care management practices. J Am Geriatr Soc 2003 May;51(5):710-714. [doi: 10.1034/j.1600-0579.2003.00218.x] [Medline: 12752849]
6. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: a novel mHealth system for supporting self-care in management of complex and chronic conditions. JMIR Mhealth Uhealth 2013 Jul 11;1(2):e10 [FREE Full text] [doi: 10.2196/mhealth.2391] [Medline: 25100682]
7. HIPAA Home. Health and Human Services. URL: https://www.hhs.gov/hipaa/index.html [accessed 2021-06-29]
8. Bendixen RM, Fairman AD, Karavolis M, Sullivan C, Parmanto B. A user-centered approach: understanding client and caregiver needs and preferences in the development of mHealth apps for self-management. JMIR Mhealth Uhealth 2017 Sep 26;5(9):e141 [FREE Full text] [doi: 10.2196/mhealth.7136] [Medline: 28951378]
9. Fairman AD, Dicianno BE, Datt N, Garver A, Parmanto B, McCue M. Outcomes of clinicians, caregivers, family members and adults with spina bifida regarding receptivity to use of the iMHere mHealth solution to promote wellness. Int J Telerehabil 2013;5(1):3-16 [FREE Full text] [doi: 10.5195/ijt.2013.6116] [Medline: 25945209]
10. Fairman AD, Yih ET, McCoy DF, Lopresti EF, McCue MP, Parmanto B, et al. Iterative design and usability testing of the iMHere system for managing chronic conditions and disability. Int J Telerehabil 2016;8(1):11-20 [FREE Full text] [doi: 10.5195/ijt.2016.6194] [Medline: 27563387]
11. Setiawan IMA, Zhou L, Alfikri Z, Saptono A, Fairman AD, Dicianno BE, et al. An adaptive mobile health system to support self-management for persons with chronic conditions and disabilities: usability and feasibility studies. JMIR Form Res 2019 Apr 25;3(2):e12982 [FREE Full text] [doi: 10.2196/12982] [Medline: 31021324]
12. Zhou L, DeAlmeida D, Parmanto B. Applying a user-centered approach to building a mobile personal health record app: development and usability study. JMIR Mhealth Uhealth 2019 Jul 05;7(7):e13194 [FREE Full text] [doi: 10.2196/13194] [Medline: 31278732]

13. Zhou L, Saptono A, Setiawan IMA, Parmanto B. Making self-management mobile health apps accessible to people with disabilities: qualitative single-subject study. JMIR Mhealth Uhealth 2020 Jan 03;8(1):e15060 [FREE Full text] [doi: 10.2196/15060] [Medline: 31899453]
14. Yu D, Parmanto B, Dicianno B. An mHealth app for users with dexterity impairments: accessibility study. JMIR Mhealth Uhealth 2019 Jan 08;7(1):e202 [FREE Full text] [doi: 10.2196/mhealth.9931] [Medline: 30622096]
15. Yu DX, Parmanto B, Dicianno BE, Pramana G. Accessibility of mHealth self-care apps for individuals with spina bifida. Perspect Health Inf Manag 2015;12:1h [FREE Full text] [Medline: 26755902]
16. Yu DX, Parmanto B, Dicianno BE, Watzlaf VJ, Seelman KD. Accessibility needs and challenges of a mHealth system for patients with dexterity impairments. Disabil Rehabil Assist Technol 2017 Jan;12(1):56-64. [doi: 10.3109/17483107.2015.1063171] [Medline: 26153097]
17. Dicianno BE, Fairman AD, McCue M, Parmanto B, Yih E, McCoy A, et al. Feasibility of using mobile health to promote self-management in spina bifida. Am J Phys Med Rehabil 2016 Jun;95(6):425-437. [doi: 10.1097/PHM.0000000000000400] [Medline: 26488144]
18. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE. Development of mHealth system for supporting self-management and remote consultation of skincare. BMC Med Inform Decis Mak 2015 Dec 30;15:114 [FREE Full text] [doi: 10.1186/s12911-015-0237-4] [Medline: 26714452]
19. Kryger MA, Crytzer TM, Fairman A, Quinby EJ, Karavolis M, Pramana G, et al. The effect of the interactive mobile health and rehabilitation system on health and psychosocial outcomes in spinal cord injury: randomized controlled trial. J Med Internet Res 2019 Aug 28;21(8):e14305 [FREE Full text] [doi: 10.2196/14305] [Medline: 31464189]
20. 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]
21. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
22. Sockolow PS, Weiner JP, Bowles KH, Abbott P, Lehmann HP. Advice for decision makers based on an electronic health record evaluation at a program for all-inclusive care for elders site. Appl Clin Inform 2011;2(1):18-38 [FREE Full text] [doi: 10.4338/ACI-2010-09-RA-0055] [Medline: 23616858]
23. Fretwell MD, Old JS, Zwan K, Simhadri K. The Elderhaus Program of All-inclusive Care for the Elderly in North Carolina: improving functional outcomes and reducing cost of care: preliminary data. J Am Geriatr Soc 2015 Mar;63(3):578-583. [doi: 10.1111/jgs.13249] [Medline: 25752225]
24. Stefanacci RG, Reich S, Casiano A. Application of PACE principles for population health management of frail older adults. Popul Health Manag 2015 Oct;18(5):367-372. [doi: 10.1089/pop.2014.0096] [Medline: 25647691]

Abbreviations

CLASS: Community Living and Support Services
CPP: Community Partners Program
ICE: In Case of Emergency
mHealth: mobile health
PHR: Personal Health Record

Edited by A Kushniruk; submitted 30.06.21; peer-reviewed by 찬 장, M Salimi; comments to author 27.08.21; revised version received 21.10.21; accepted 11.11.21; published 04.03.22.

Please cite as:
Ward RE, Setiawan IMA, Quinby E, Fair M, Ambadar Z, Parmanto B, Dicianno BE
Mobile Health to Support Community-Integration of Individuals With Disabilities Using iMHere 2.0: Focus Group Study
JMIR Hum Factors 2022;9(1):e31376
URL: https://humanfactors.jmir.org/2022/1/e31376
doi:10.2196/31376
PMID:35254274

Original Paper

Lessons for Uptake and Engagement of a Smartphone App (SURE Recovery) for People in Recovery From Alcohol and Other Drug Problems: Interview Study of App Users

Joanne Neale[1], BA, MA, CQSW, DPhil; Alice May Bowen[1], BSc, MSc

National Addiction Centre, King's College London, London, United Kingdom

Corresponding Author:
Alice May Bowen, BSc, MSc
National Addiction Centre
King's College London
4 Windsor Walk
Denmark Hill
London, SE5 8BB
United Kingdom
Phone: 44 7511319855
Email: alice.m.bowen@kcl.ac.uk

Abstract

Background: Mobile health apps promoting health and well-being have substantial potential but low uptake and engagement. Barriers common to addiction treatment app uptake and engagement include poor access to mobile technology, Wi-Fi, or mobile data, plus low motivation among non–treatment-seeking users to cut down or quit. Working with people who used substances, we had previously designed and published an app to support recovery from alcohol and other drug problems. The app, which is available for free from the Apple App Store and Google Play, is called *SURE Recovery.*

Objective: The aim of this paper is to undertake a qualitative study to ascertain end users' views and experiences of the SURE Recovery app, including how it might be improved, and present the findings on uptake and engagement to assist other researchers and app developers working on similar apps for people experiencing alcohol and other drug problems.

Methods: Semistructured telephone interviews were conducted with 20 people (n=12, 60%, men and n=8, 40%, women aged 25-63 years; all identifying as White British) who had varied patterns of using the app. The audio recordings were transcribed, and the data were coded and analyzed through Iterative Categorization.

Results: Analyses identified three main factors relevant to *uptake* (discoverability of the app, personal relevance, and expectations and motivations) and three main factors relevant to *engagement* (the appeal and relevance of specific features, perceived benefits, and the need for improvements). The findings on uptake and engagement were largely consistent with our own earlier developmental work and with other published literature. However, we additionally found that uptake was strongly affected by first impressions, including trust and personal recommendations; that users were attracted to the app by their need for support and curiosity but had relatively modest expectations; that engagement increased if the app made users feel positive; and that people were unlikely to download, or engage with, the app if they could not relate to, or identify with, aspects of its content.

Conclusions: Incorporating end-user views into app design and having a network of supportive partners (ie, credible organizations and individuals who will champion the app) seem to increase uptake and engagement among people experiencing alcohol and other drug problems. Although better digital literacy and access to devices and mobile data are needed if addiction recovery apps are to reach their full potential, we should not evaluate them based only on observable changes in substance use behaviors. How using an app makes a person feel is more transient and difficult to quantify but also relevant to uptake and engagement.

(JMIR Hum Factors 2022;9(1):e33038) doi:10.2196/33038

KEYWORDS

apps; digital literacy; mHealth; substance use; recovery; qualitative; mobile phone

Introduction

Background

The use of mobile health (mHealth) apps is increasingly common among clinical and nonclinical populations, including people who use substances. With the ownership of smartphones, tablets, and wearable devices growing and access to wireless networks expanding, the number of apps relating to substance use has proliferated and downloads of the most popular apps have risen [1,2]. Alongside their potential reach, mHealth apps are convenient (they can be accessed anytime and anywhere), are low cost, and can overcome some of the barriers to accessing standard treatment (such as strict appointment times, lengthy distances to travel, concerns regarding childcare, and stigma) [2-5]. Apps can also facilitate personalized (*tailored*) support; offer opportunities for real-time relapse prevention, treatment, and aftercare [6-8]; and are accessible to people already in treatment as well as those not currently accessing services [7].

Addiction-related apps commonly include blood alcohol content calculators, service finders, other information and resources, games to distract from cravings, strategies to increase motivation, functions to enhance social support, and tools to monitor progress [1,2,8]. Accordingly, they offer valuable opportunities for diagnostics, measurement, treatment, and recovery [1]. Evidence suggests that users particularly like behavior tracking and remote access to advice and information [8]. In addition, they appreciate the portability of apps, their discretion (given the stigma around addiction), and the fact that they tend to be free or low cost [2,8]. However, alcohol-related apps are often designed for entertainment or to promote, rather than reduce, alcohol use [7], and cannabis apps tend to be for informational or recreational purposes [9]. Furthermore, treatment-oriented apps predominantly focus on tobacco and alcohol, rather than illicit drugs, and have mostly been developed in the United States, potentially limiting their relevance internationally [8-11].

Importantly, apps promoting health and well-being also tend to suffer from low uptake and engagement [11]. Uptake is the act of downloading and installing a smartphone app, whereas engagement refers to both the extent (eg, amount, frequency, duration, and depth) of use and the user's personal experience as characterized by, for example, their attention, interest, or mood [11,12]. Barriers to addiction treatment app uptake and engagement include poor access to mobile technology, Wi-Fi, or mobile data [13,14] as well as low motivation among non–treatment-seeking users to cut down or quit [9]. Compounding these limitations, evidence for the effectiveness of smartphone apps in addressing alcohol and other drug problems is weak [2,7,8,15]. This may be related to the poor quality of much app content; for example, apps rarely include empirically based behavior change techniques (other than self-monitoring) [2,9,15,16] and app developers often lack personal or clinical experience of addiction [8].

Before developing a new recovery app to help people reduce or cease their alcohol and other drug use and improve their quality of life, it seems sensible to ask whether there is a genuine demand. This question is likely to be especially important if people who use substances do not want to change their behavior, have limited resources or complex needs, and experience high levels of digital exclusion. Although a candid answer may be that demand is likely to be weak, it still seems wrong to perpetuate inequalities and lack of choice by failing to offer digital options for those who might be interested or benefit. A better alternative would be to work with the target population to develop an app that might support at least some people in addressing their substance use while also seeking to learn from the process and results. It is on this basis that we developed the SURE Recovery app [17].

People with personal experience of addiction had asked us to convert our two validated pen and paper measures, the Substance Use Recovery Evaluator (*SURE*; a 21-item measure of addiction recovery) [18] and the Substance Use Sleep Scale (*SUSS*; a 23-item measure of sleep problems experienced by people using substances) [19] into an app that they could complete on their mobile phones and tablet computers. They explained that they wanted to record and refer back to their SURE and SUSS scores, and they also expressed a desire for personalized feedback. Further discussions suggested that they would like to see the 2 measures supplemented with other features that might promote recovery from alcohol and other drug problems. Both SURE and SUSS had been developed collaboratively with people who had experience of addiction, and we continued this joint working by adopting a co-design approach when developing the SURE Recovery app.

Co-design involves *end users* throughout the design process as active partners [20], providing people whose lives might be affected by a problem with a voice in its solution [21,22]. Evidence suggests that the inclusion of end users in the early stages of the design process leads to better outcomes and more benefits compared with ideas developed by designers alone [23]. Our co-design approach was completed following the Double Diamond design process, which is a framework widely used in the design industry. This involves four distinct phases—(1) discover, (2) define, (3) develop, and (4) deliver—that are repeated in iterative cycles to ensure that end-user feedback is incorporated throughout [24]. To this end, we conducted interviews, focus groups, review meetings, and testing sessions with nearly 50 people in recovery or actively using substances. In addition, our team comprised people with personal experience of alcohol and other drug problems, clinicians (addiction psychologists and psychiatrists), and academics (social scientists and statisticians).

During the *discover* phase of our work, interview and focus group participants explained that they valued different types of formal and informal support, enjoyed connecting with others in similar situations, appreciated being busy and distracted, felt that keeping a log of their recovery was helpful, and wanted advice on the types of support available. When asked to comment on the design and content of other apps, they expressed preference for a clear layout, bright colors, simplicity, tracking features, inspirational quotations, nonjudgmental and supportive language, an opportunity to share artwork, and the ability to connect with others. In contrast, they disliked apps that seemed *busy* or *crowded*, had too much text, contained advertisements, or looked *technical*. From this feedback, a long list of potential

app features was created. This included information and advice; a directory of services; opportunities to meet or share personal stories, experiences, advice, and artwork; tracking (progress, mood, or problems); a way to be reminded of the app; encouragement and motivation; and sleep tracking. Following discussions during the *define* phase, the team narrowed the options down to 6 features that were viable within the project budget and time frame, plus a set of optional research questions covering basic demographics, substance use, and treatment-related topics.

App Features

The six features included in the SURE Recovery app are as follows:

1. A recovery tracker (this allows people to monitor their own recovery through SURE and receive personalized feedback and a score that can be viewed on a graph)
2. A sleep tracker (this works in a similar way to the recovery tracker, enabling people to monitor their sleep through SUSS and receive personalized feedback, a score, and a graph)
3. Artwork (app users can submit their artwork for potential display in the banner of the app home screen)
4. Diary (a private space where people can record their thoughts and feelings)
5. Naloxone (an instructional video on how to use the life-saving medication naloxone in the event of an opioid overdose, plus informational resources and a knowledge tracker to measure overdose management competency)
6. Resources (free access to a book,*The Everyday Lives of Recovering Heroin Users*, which is based on the lived experiences of people in recovery) [25]

After much reflection, we did not include a social feature, where app users could *chat* and share experiences and advice, because the team did not have adequate resources to monitor the chat in a way that would ensure app user safety at all times.

The SURE Recovery app has been available to download for free from the Apple App Store and Google Play since October 2019 (>2200 downloads by May 31, 2021). It was updated with a temporary COVID-19 pop-up feature (comprising COVID-19 resources, information, and a new research question) in April 2020. In March 2021, the temporary COVID-19 feature was replaced by a more permanent and dynamic *hot topic of the month* feature (allowing information on a contemporary relevant issue to be displayed, an associated research question to be asked, and key findings from any responses received to be posted back into the body of the app). Given the importance of understanding what end users thought of SURE Recovery, including their views on whether and how it might be improved going forward, 2 members of the SURE App team (JN and AMB) also conducted a qualitative study. The aim of this paper is to present our findings on uptake and engagement to assist others developing similar apps for people experiencing alcohol and other drug problems.

Methods

Ethics Approval

Ethical approval for the qualitative study was received from the research ethics committee of King's College London (HR-19/20-17338).

Overview

Data were generated through semistructured telephone interviews conducted with 20 people who had downloaded the SURE Recovery app. When signing up for the app, all users are provided with a link to a web-based information sheet and asked if they are willing to share their anonymized data for quantitative research. If they agree, they are given the option to consent within the app. App users are next asked if they would be willing to be contacted by a researcher to participate in further research relating to SURE Recovery. Those who agreed to both share their data and be contacted for further research (N=620) were entered into a pool of potential participants for the qualitative study. For pragmatic reasons (the cost of international telephone calls, time differences, increased likelihood of poor telephone reception, and language differences), of the 620 respondents, we excluded 241 (38.9%) who were based outside the United Kingdom, leaving 379 (61.1%) potential participants. From these, we sampled purposively to include people who had used the app once or twice only, occasionally, and frequently. As a secondary strategy, and to be as inclusive of views as possible, we also endeavored to sample people with a mix of demographic, substance use, and treatment characteristics.

Author AMB first contacted potential participants through the email address they had used when registering for the app. In total, 107 app users were contacted in this way over the course of 10 months (May 2020 to February 2021). The email sent contained basic information regarding the qualitative study and invited the recipient to respond with a telephone number if they wanted to hear more. A maximum of 3 invitation emails were sent to each person. People who responded positively (24/107, 22.4%) were then emailed the study information sheet and consent form and asked to select a time when they could be interviewed. An additional telephone call was offered to anyone who wanted to know more about the study before deciding whether to participate. At this point, of the 24 participants, 4 (17%) withdrew their interest, whereas 20 (83%) agreed to continue. AMB conducted all interviews by telephone, securing verbal consent before each interview started. Although the target number of participants had been 30, recruitment ceased after 107 app users had been contacted and 20 interviews had been completed. This was because both authors believed that data saturation had been achieved: comments regarding the app from new interviews were largely repeating comments from earlier interviews and no new themes or topics seemed to be emerging [26].

All interviews were audio recorded and followed a topic guide that covered the participant's background (general life circumstances, health, education, employment, substance use, and treatment history), initiation to SURE Recovery (how the participant had first heard about SURE Recovery, their expectations, motivations, goals, and reasons for downloading

the app), use of SURE Recovery (frequency; duration; cessation; when, where, and how the app was used; barriers to use; and features most used), positive views of SURE Recovery (features liked and any benefits of use), negative views of SURE Recovery (features disliked and any negative consequences of use), and potential improvements to SURE Recovery (suggested improvements, strategies for overcoming barriers to use, and ideas for new features). Interviews lasted 18-73 minutes, and participants were paid £20 (US $27.20) as compensation for their time.

Data Analyses

Data analyses followed the stages of Iterative Categorization [27,28]. To begin with, the audio files were transcribed verbatim by a professional transcription service and the transcriptions were uploaded to the software data management program MAXQDA (version 2018.2; VERBI Software GmbH) [29]. Next, both authors jointly devised a simple coding frame that mirrored the interview topic guide. Subsequently, AMB indexed all transcribed text to one or more of the codes and exported the indexed data from the software program into Microsoft Word documents (1 Word document per code). Each Word document was then reviewed line by line (either by AMB or by JN) to identify patterns and themes in the data. To this end, all indexed text was summarized into bullet points, and the bullet points were iteratively grouped into themes and categories that were in turn summarized (1 summary per code). Next, JN combined the summaries from each code into 1 main findings document that provided an overarching descriptive account of the participants' views. JN then systematically reviewed the main findings document for material relating to SURE Recovery uptake and engagement before AMB checked and confirmed the findings.

Results

Participant Characteristics

Table 1 presents basic data relating to all people who downloaded the app between October 1, 2019, and May 31, 2021; consented to share their data; and consented to be contacted for further research (N=620); and all people who downloaded the app between October 1, 2019, and May 31, 2021; consented to share their data; consented to be contacted for further research; and were based in the United Kingdom (379/620, 61.1%); as well as all participants who contributed a qualitative interview (20/379, 5.3%). These data are provided to contextualize the qualitative study participants within the wider body of app users. Table 1 suggests that those participating in the qualitative interviews may have been more likely to have ever had a problem with use of opioids or alcohol, to have attended mutual aid meetings or peer support groups in the last week, and to be in paid work than other app users. In contrast, they were potentially less likely to have used substances or to have been in formal treatment in the last week. This might simply reflect the fact that individuals who are more stable in recovery are more willing and able to participate in a qualitative telephone interview than those who are still regularly using substances and in formal treatment.

As seen in Table 1, of the 20 people who participated in a qualitative interview, 12 (60%) were men and 8 (40%) were women. They had a mean age of 43 (range 25-63) years, and all were White British. The qualitative interviews provided additional and more comprehensive demographic information and drug use data about the study participants. When interviewed, 8 (40%) said that they were in paid employment (of these 8 participants, 5, 63%, said that they worked in the drug treatment sector); 10 (50%) reported that their substance of choice was alcohol and 10 (50%) said that their substance of choice was another psychoactive drug; and 8 (40%) had ever injected a drug. Although all 20 (100%) participants identified as ever having had a problem with alcohol or other drugs, 13 (65%) said that they had not used any substances in the last month and 6 (30%) said that they were neither currently receiving formal treatment nor attending any mutual aid or peer support groups.

Table 1. SURE Recovery app user characteristics (October 1, 2019, to May 31, 2021).

Characteristics	All users who consented to share their data and be contacted for further research (N=620)	All users based in the United Kingdom who consented to share their data and be contacted for further research (n=379)	Users participating in a qualitative interview (n=20)
Gender, n (%)			
Male	269 (43.4)	192 (50.7)	12 (60)
Female	333 (53.7)	182 (48)	8 (40)
Other	7 (1.1)	1 (0.3)	0 (0)
Prefer not to say	11 (1.8)	4 (1)	0 (0)
Age (years)			
Value, mean (SD)	41 (11)	42 (10.7)	43 (10.5)
Missing, n (%)	82 (13.2)	32 (8.4)	0 (0)
Ethnicity[a] (White British), n (%)	N/A[b]	N/A	20 (100)
App users who completed optional questions on first-ever use of the app[c], n (%)	308 (49.7)	184 (48.5)	6 (30)
Participated in paid employment during the last week, n (%)			
Yes	176 (57.1)	110 (59.8)	5 (83.3)
No	132 (42.9)	74 (40.2)	1 (16.7)
Ever had a problem with use of heroin or other opiates, n (%)			
Yes	89 (28.9)	43 (23.4)	3 (50)
No	219 (71.1)	141 (76.6)	3 (50)
Ever had a problem with use of alcohol, n (%)			
Yes	203 (65.9)	136 (73.9)	5 (83.3)
No	105 (34.1)	48 (26.1)	1 (16.7)
Any substance use in the last week, n (%)			
Yes	229 (74.4)	129 (70.1)	1 (16.7)
No	79 (25.6)	55 (29.9)	5 (83.3)
Contact with community drug and alcohol treatment services in the last week, n (%)			
Yes	103 (33.4)	65 (35.3)	1 (16.7)
No	205 (66.6)	119 (64.7)	5 (83.3)
Attended mutual aid meetings or a peer support group in the last week, n (%)			
Yes	114 (37)	72 (39.1)	3 (50)
No	194 (63)	112 (60.9)	3 (50)
In residential treatment during the last week, n (%)			
Yes	15 (4.9)	8 (4.3)	0 (0)
No	293 (95.1)	176 (95.7)	6 (100)

[a]Apple does not permit developers to require personal information that is not directly relevant to the app's core functionality at registration. We decided to not include an optional ethnicity question, given the number of sensitive optional questions regarding substance use already being asked and concerns that many potential users may consider an ethnicity question irrelevant or be frustrated by a long scroll list that may make finding their own ethnicity difficult. The lack of ethnicity data has resulted in a limitation in our analyses, which we discuss further in the *Limitations* section.

[b]N/A: not applicable.

[c]SURE Recovery users can return and complete optional questions at any time when using the app. For consistency, only data entered by users on their first occasion of using the app are reported. This means that the number of responses to various questions in the table is less than the number of app users (n=308, n=184, and n=6 rather than N=620, n=379, and n=20).

Participants' Use of SURE Recovery

Consistent with our recruitment strategy, use of the app by our qualitative study participants varied greatly. Thus, we interviewed people who had recently downloaded the app but had not yet started to use it; had used it once or twice and then stopped; had used it frequently initially but were now using it less; were using it daily; and were using it occasionally. Further details regarding the participants' use of SURE Recovery are shown in Table 2.

Turning to our main qualitative analyses, we identified 3 key factors relevant to decisions to download and install the app, that is, *uptake*. These were (1) discoverability of the app, (2) personal relevance, and (3) expectations and motivations. In addition, we found that 3 key factors affected use and experiences, that is, *engagement*. These were (1) the appeal and relevance of specific features, (2) perceived benefits, and (3) the need for improvements. We next present our findings using pseudonymized quotations to illustrate salient points.

Table 2. Participants' use of SURE Recovery (N=20)[a].

Type of SURE Recovery use	Values, n (%)
Current use	
Yes	11 (55)
No	6 (30)
About to start or restart	3 (15)
Frequency of current use	
Daily	2 (10)
Weekly	6 (30)
Monthly	3 (15)
Unknown	9 (45)
Data transfer source	
Wi-Fi	9 (45)
Cellular	2 (10)
Wi-Fi and cellular	8 (40)
Unknown	1 (5)
Device used	
Android phone	6 (30)
iPhone	6 (30)
iPhone and Android tablet	1 (5)
Unknown	7 (35)

[a]Comparable data are not available for app users who did not participate in a qualitative interview.

Uptake

Discoverability of the App

Knowing how participants first heard about SURE Recovery provides potentially important information on how addiction-related apps are *discovered* and thus how they might be introduced or even *advertised* to potential users. Most study participants had first learned about SURE Recovery either through a key worker, support worker, or professional who was working with them or from browsing or searching for addiction-related information and support on the web. A few participants who were employed as recovery workers within the addiction treatment sector said that they had come across the app during the course of their work. In addition, 1 (5%) had read about the app in a newsletter and 2 (10%) had been introduced to it by a friend or peer in recovery:

> *So, one of the girls that is in NA [Narcotics Anonymous] with me, she was telling me about it [the app]. Because I was saying I was struggling with my sleep...And she said to go on this app and it'll like help you with your sleep.* [Daisy, female, aged 39 years, weekly app user]

Participants who had been introduced to SURE Recovery by a key worker, support worker, or professional explained that the worker had directed them to the app using a hyperlink sent in an email, by signposting them to a website, or by sending information in hard copy through the post. All participants had then successfully downloaded the app themselves. Participants who had found SURE Recovery by browsing or searching on the web mostly said that they had been trawling the webpages of a recovery organization to look for support for themselves or for others, although 1 (5%) had noticed it on social media (Instagram), and another had found it through the Apple App Store. Of the 4 participants who had learned of the app through

their employment, 2 (50%) said that they had been proactively researching apps and 2 (50%) explained that details had been cascaded down to them from managers:

> *Part of the team I lead on is around sort of providing psychosocial interventions and group work, and...it [SURE Recovery] came up as one of the potential tools that we use with our client group.* [Frank, male, aged 43 years, tried app a few times]

Several participants clarified that a key factor prompting them to seek out a recovery app was the COVID-19 pandemic because this had created problems physically attending services. Some participants said that they had considered and downloaded several apps before settling on SURE Recovery, and a key reason for choosing SURE Recovery was that they thought that they might have heard of either the SURE measure or the SURE Recovery app already.

Personal Relevance

Although SURE Recovery had been developed for anyone in recovery or thinking about recovery from alcohol and other drug problems, our participants identified subgroups of people for whom they thought the app would be more or less appropriate. Most frequently, they suggested that the app would be more suitable for people who were in *early recovery* rather than for those who had been abstinent or stable for some time. The main reason they gave for this was that people who were in long-term recovery would be more likely to score consistently well on the SURE measure, meaning that they had little scope for improvement and therefore little incentive to return to the app to complete the measure again. In contrast, they said that someone at the start of their recovery journey would be able to complete SURE over time and see rewarding changes as their SURE scores increased on the graph:

> *If I was speaking to somebody that I realised was in the contemplation phase, about to begin recovery, I would say, "Look, here's this SURE app. Jump on this, put your score in. I guarantee in a month's time you will see that it's having tangible benefits." Because it's a great way to actually show yourself and remind yourself that you're in recovery for a reason.* [Ben, male, aged 46 years, monthly app user]

In addition, many participants thought that SURE Recovery would be more useful for people who had a problem with drugs, particularly heroin, rather than alcohol. This, they explained, was because the balance of the app content seemed to be on opioids, with 33% (2/6) of the features (the naloxone feature and the reading section) being very specific to heroin. Several participants stated that this made the app feel less relevant to them personally:

> *Well, it's just obviously for heroin addicts, isn't it? But that might be just because I've never took heroin, and I just don't relate to it. I don't know...Obviously people that have used heroin, it's probably for them.* [Laura, female, aged 35 years, daily app user]

A small number of participants added that the emphasis on heroin was *off-putting*, and 1 (5%) reported that questions within the app on homelessness and being in prison made them doubt whether the app was really for them because they did not identify with these issues.

More generally, several participants stated that the app would be particularly relevant to people who were concerned about privacy and to those who did not like mutual aid meetings. In this regard, participants emphasized that the app provided a nonjudgmental and safe space for people to find different types of information and support without having to share their personal data. Of 20 participants, 1 (5%) added that the app could be useful for people who did not have much external support. Others felt that specific features (particularly the artwork feature, sleep tracker, or diary) might interest some individuals, and several participants had forwarded information regarding the app to peers who might (they thought) appreciate these functions.

Notwithstanding these opportunities, participants also argued that the app might be less helpful for people who were not comfortable with technology, did not have a smartphone, did not have access to mobile data, were homeless, or were using substances very heavily. In addition, some questioned whether people who were not ready to address their addiction would be interested in using the app or whether someone who was having a difficult time would be willing to engage with it:

> *For me, being scored [using the tracker features] works. For other people, if they go in and out of lapses, or even a full relapse, they may not want the added pressure of their scores getting worse. Because that may then...lead to further use.* [Ben, male, aged 46 years, monthly app user]

Expectations and Motivations

Most participants reported that they did not have any, or any particular, expectations before downloading SURE Recovery. Some explained that they just thought they would *give it a go* and hoped that it would offer them help or something to assist them in staying abstinent or sober. Others clarified that they did not have any *big* expectations and were simply curious. As 1 (5%) of the participants explained, addiction is *complicated* and cannot be *cured by an app*, although it is *an additional tool*. Less positively, another participant stated that they were not convinced that the app would be of much use because they preferred face-to-face meetings but had been urged to try it by their drug worker.

At the time when they downloaded SURE Recovery, some participants said that they were feeling positive and wanted to use the app to stay focused and maintain progress. In contrast, others said that they were *struggling*, *not feeling great*, or *in a terrible state* and therefore were seeking new forms of support. For example, 2 (10%) said that the app interested them because it might offer something different from Alcoholics Anonymous and Narcotics Anonymous, and another explained that they were attracted to the app because they did not find it easy to talk to people. Other participants commented that they had liked the look of the app because it seemed *easy to understand*, *simple to use*, and *always there*:

> *It's something that's there at the time, you know, when you're having your thoughts, rather than, "Oh, you*

> *know, I've got to go and make an appointment to see...my counsellor"...It's immediately there...and that's what you need.* [Amy, female, aged 43 years, weekly app user]

In addition, a number of participants referred to particular features or content that had piqued their interest. Most often this was the recovery-tracking feature, which they said would enable them to monitor and reflect on their recovery journeys. However, others explained that they had been drawn by the sleep feature, and some mentioned the diary because this offered them somewhere to write down their thoughts, feelings, and activities. Several participants also reported that they had been attracted to the app because they thought that it might be able to help them with general recovery goals such as maintaining sobriety, avoiding relapse, engaging in self-help, and taking responsibility for themselves.

First impressions additionally seemed particularly important. In this regard, several participants stated that the app had seemed *different from other apps*, offered a range of content, looked as though it might provide something new, seemed to have been well researched, and was not simply about *counting* days sober. Others confirmed that it looked interesting and useful (although 1 (5%) of the participants said that they had been a little concerned that it would be too complicated for them). Some also reported that they had seen a positive review on the Apple App Store or felt that the app was *trustworthy* because it had been recommended to them by someone they respected or had been developed by a university:

> *You can tell from the App Store that it was developed by [name of university], and you know, like there was research being put into it. So...I think I just trusted it a bit more.* [Lucy, female, aged 28 years, tried app a few times]

Engagement

The Appeal and Relevance of Specific Features

When participants discussed how and why they continued to engage with the app (or why they disengaged from it), the appeal and relevance of specific features were central. Most of them said that they used the recovery tracker more than any other feature. Generally, participants thought that it was motivating, interesting, useful, or fun to track their scores. Despite this, a small number of participants had not noticed the recovery tracker, and 1 (5%) had dismissed it as being too much effort to complete. In addition, a few participants thought that it was not relevant to them. This, they said, was because they scored high initially and therefore felt that they had no way of progressing, they received the same scores each time they completed it and therefore lost interest, or they thought that the questions did not apply to people such as themselves who were at a more established stage of recovery:

> *A lot of the questions were loaded towards like stable housing, and I think in recovery your outlook changes to the fact that what you need's much more than that. And it [recovery tracker] didn't go deep enough for me.* [Luke, male, aged 44 years, tried app a few times]

Although many participants had used the sleep tracker, some stated that they did not use it because they slept well, slept badly, or accessed other apps for sleep monitoring. The artwork feature was, meanwhile, generally appreciated, with participants variously describing it as *interesting*, *cool*, *brilliant*, and a *nice touch*. However, a few participants found it confusing and said that they did not see how it linked to the rest of the app or how people might use it if they did not have artwork to submit. The diary feature was used regularly and received some of the most positive feedback, with participants stating that they enjoyed recording their feelings (and, to a lesser extent, activities) and then looking back over their entries. Nonetheless, a few participants said that they had not used the diary feature because they did not keep a diary, preferred to record things on paper, or feared that their entries might be read by someone else, particularly if they lost their phone:

> *So, I suppose the aspect of the diary is [that] I would just worry somehow if I lost my phone...if somehow what you're writing...they [diary entries] are personal and private to you.* [Lucy, female, aged 28 years, tried app a few times]

In contrast, the naloxone feature had not been widely used and generated quite mixed responses. A few participants appreciated having information regarding naloxone and overdosing within the app and stated that this could address misconceptions regarding overdose or would be helpful if someone witnessing an overdose panicked and forgot what to do. Nonetheless, others felt that this component of the app was not relevant to them because they were in long-term recovery or had never used heroin. The reading feature similarly evoked mixed views. Several participants said that the *Everyday Lives of Recovering Heroin Users* book did not interest them, and a participant complained that it was too long, whereas others said that they were enjoying reading it:

> *I just like it [Everyday Lives of Recovering Heroin Users book]...I like it just that it's personal, it's personal stories, it's true, you know. It relates to obviously my life and things.* [Claire, female, aged 49 years, daily app user]

Most participants said that they had completed some of the research questions, with 1 (5%) emphasizing how important it is to share views and experiences with researchers to help others. Meanwhile, only a small number of participants reported that they liked the temporary COVID-19 feature, with others stating that they were tired of hearing about COVID-19 and therefore not interested in engaging with this content.

Perceived Benefits

Participants identified both practical and emotional benefits from using SURE Recovery, which seemed likely to maintain their interest and engagement. These benefits were reflected in both how and when people used the app. For example, some said that they used SURE Recovery when they were feeling relaxed to reflect back on their day or to help reinforce positive emotions, whereas others said that they used it when they believed that their mind might wander to drugs, were feeling concerned about their substance use, were feeling bored and

needed a distraction, or thought that they might experience cravings:

> *I use it generally quite late at night. And I think it's because that's when my mind goes wandering to my sort of craving.* [Liam, male, aged 33 years, weekly app user]

Both practical and emotional benefits were also evident when participants discussed why they liked using particular app content and features. Thus, participants reported that the recovery tracker was useful because it enabled them to look back over their scores and see their progress, identify changes they wanted to make, and receive advice on how to advance their recovery. In addition, some stated that the feedback incentivized them to *keep going*, directed them to useful resources, was uplifting, and *gave them a boost* on a bad day. Moreover, they *enjoyed* completing the questions. One participant (5%) added that the sleep measure facilitated discussions with their physician regarding sleep, whereas others appreciated the naloxone feature because they said that it provided important information on how to save a life and made them feel more confident about responding to an overdose if needed:

> *With the naloxone, if somebody goes over [overdoses], it's there ready, you know. I mean I've had training on naloxone, but...nobody knows how [they are] going to react when it [an overdose] happens. It's just nice...that there's something in your back pocket.* [James, male, aged 54 years, weekly app user]

Several participants additionally stated that the diary feature was valuable because it allowed them to empty their heads and put all their thoughts down in one place and it could be used as a *gratitude journal* (that is, a place to record and reflect on things for which they were thankful). Some enthused about the artwork feature and explained how this inspired them and lifted their mood, whereas others said that reading the book and learning about the experiences of others in recovery was enjoyable and could help people feel less alone. In addition, some said that the embedded links to external websites provided helpful information and a route to additional forms of assistance.

More generally, participants reported that the app was useful because it could be accessed at any time or in any place and, for some, this seemed a better option than visiting a therapist, who would need an appointment. Participants also confirmed that the app was simple to understand, did not use up much mobile data, and felt *friendly toward people who had experience of addiction*, which made them feel that they could be honest when entering their data. Equally, participants said that they appreciated the variety of content and functions and noted how not using the app for a while might alert them to an impending relapse:

> *And I think that's something that's good with the app, because you can sort of measure like the time in between using them [tracker features]. You...might look at it and think, "Oh, I haven't been on there for ten days. Something's not right.*" [Charlotte, female, aged 43 years, daily app user]

In terms of concrete benefits, several participants said that the app had brought stability to their lives and had supported them to remain stable or abstinent. However, most said that their behavior had not changed as a direct result of using the app, although a few noted that engaging with the app had been part of wider positive behavior change that they had made in their recovery. Significantly, none of the participants identified any reason why the app might be unhelpful or harmful, although 1 (5%) cautioned that it was not a replacement for other forms of support and people would likely need additional help, particularly in early recovery. A few participants also felt that there were too many questions and the app was not participative enough to be helpful as an *intervention*.

The Need for Improvements

Overall, there was no suggestion that the app needed to be improved in terms of usability, although a small number of participants felt that the language within the app could be simplified. Several participants also reported that they were uncertain who exactly the app was for and thought that it might be better to have a single target audience, such as people in early recovery. In addition, some participants felt that engagement with the app might increase if it had notifications and reminder features so that people would remember to complete the measures and diary each day:

> *Reminders as well, you know, daily reminders for people, are quite important..."What have you done today for your recovery?" That kind of stuff. I think that stuff's pretty...important.* [Luke, male, aged 44 years, tried app a few times]

In terms of specific features, various participants suggested that the recovery feature could be improved by having more questions for people who were further along in their recovery, scope for scoring higher, and additional feedback on how to improve their recovery score. Other participants said that they would have liked more feedback on how to improve their sleep score and felt that the inclusion of meditation and relaxation aids would be useful additions. No particular improvements were suggested to the artwork feature or diary content, other than a passcode to increase the diary's security and privacy. Meanwhile, participants who thought that the app was too opiate focused expressed a desire for more reading and resources on other substances, such as a book on recovery from alcohol problems.

Turning to new features, many participants wanted to see a simple sobriety tracker that recorded an individual's number of days abstinent, whereas others recommended the inclusion of affirmations (that is, positive statements that can help people overcome negative and self-sabotaging thoughts). Participants additionally suggested including a section on other elements of well-being, such as nutrition, exercise, and mental health or mood. Finally, some thought that the app would be better if it enabled them to connect with, and talk to, others in recovery; for example, through a live chat or newsfeed or by having opportunities to submit personal stories and experiences:

> *I feel there should be like maybe where you can sign up and you can interact with other people in recovery.*

> *So that, you know, you can kind of support each other, and also meet other people that are sober and in recovery.* [Liam, male, aged 33 years, weekly app user]

Discussion

Principal Findings

Our analyses identified 3 main factors influencing uptake of, and 3 main factors influencing engagement with, the SURE Recovery app. Importantly, however, there were similarities and overlap between the uptake and engagement factors. In terms of uptake, study participants learned about the app through various sources but seemed particularly likely to download it if it came to their attention through people, services, or social media they trusted. Uptake also seemed to increase when people did not want, or were unable, to access more formal support. In addition, SURE Recovery was deemed more suitable for people who were in early recovery and users of opioids, as well as people who had access to, and a level of understanding of, technology and a degree of motivation for recovery. Although overall expectations of SURE Recovery tended to be low, particular app features appeared to pique interest and draw users in initially.

In terms of engagement, participants particularly liked the recovery and sleep trackers and the diary. The embedded research questions were considered acceptable, but the opioid-specific material (the naloxone section and, to a lesser extent, the book) were more controversial and seemed to be associated with a degree of disengagement by some. Participants attributed a range of practical and emotional benefits to using SURE Recovery, with no reports of any harm caused. Benefits included easy access to useful information, support in maintaining stability and abstinence, reinforcement of positive behavior changes, enjoyment, increased motivation to recover, and improved mood. Despite this, participants noted that the app was not a standalone *intervention* that could *cure addiction*, and some wanted greater clarity regarding the intended audience. In addition, participants recommended a range of new features, including notifications and reminders, more content on alcohol and other substances, a simple method for counting days sober, and opportunities for real-time social interaction with other people in recovery.

In practice, our research findings replicated some of the early insights we had gained from the first *discover* stage of co-designing SURE Recovery. During this initial developmental work, interview and focus group participants had also stated that they valued different types of formal and informal support, enjoyed connecting with others in similar situations, felt that keeping a log of their recovery was helpful, wanted advice on the types of support available, liked tracking features, and desired inspirational quotations. This high level of concordance between the 2 stages of our work seems to validate our decision to adopt a user-focused design process because the end users clearly appreciated the features recommended to us by people with experience of substance use during the developmental stage. This finding differs from the conclusion of a systematic review of health and well-being smartphone app uptake and engagement conducted by Szinay et al [11] that reported that study participants who discussed a hypothetical app did not always agree with those who gave their views after actually using an app. Such inconsistency with our findings merits further investigation because the good concordance we identified clearly suggests that soliciting and incorporating end-user views into app design can improve uptake and engagement later [30].

Otherwise, many of our findings were broadly consistent with both the review by Szinay et al [11] and a range of other published literature. As reported by Szinay et al [11], we found that when people were recommended the app, uptake increased; the provision of health information, reminders, self-monitoring, positive tone, social networking, and perceived utility were linked to better engagement; and app literacy skills affected both uptake and engagement. Equally, we identified support for various items of the Mobile App Rating Scale; for example, our participants appreciated fun, interest, interactivity, usability, information quality, suitability for the target audience, and credibility [31]. In line with other studies [2,7,8], we established that people liked convenience and privacy, a nonjudgmental tone, and the opportunity to see scores and monitor personal progress. More negatively, meanwhile, our analyses confirmed that uptake and engagement were likely to be undermined by poor access to mobile technology, Wi-Fi, or mobile data, and low user motivation for behavior change [9,11,14].

Importantly, our findings also yielded newer insights. First, participants were attracted to SURE Recovery based on first impressions, including the hope that the app would offer them something useful, different, or new; it looked interesting; and it seemed trustworthy because it was developed by a university or had been recommended to them. Second, participants approached SURE Recovery with very modest expectations. Indeed, they were willing to download it based on their need for support or because they were curious. Consequently, we had no need to promote, advertise, or market SURE Recovery using promises that it would *stop addiction*, *cure cravings*, or *change lives*. Third, participants indicated that a key factor in maintaining engagement was how the app made users feel. For example, our participants said that they valued the enjoyment gained from using the app, they felt motivated and heartened when they saw their scores or looked at the artwork, they were able to clear their heads when writing things down in the diary, and they experienced a connection to others when reading the book. Fourth, our analyses highlighted the significance of relatability; thus, participants seemed unlikely to download or engage with the app if they could not relate to its content, especially if that content undermined their sense of identity (for example, by inaccurately implying that they used heroin, were homeless, or committed crimes).

Taken together, our findings and reflections have potential relevance for other researchers and app developers. For example, we believe that our co-design process was critical in ensuring that SURE Recovery is user friendly, easy to understand, motivating, and trustworthy, and we recommend this collaborative way of working to others [8,30]. Nonetheless, we developed SURE Recovery in response to user demand with limited consideration of our precise target audience and whether we should be including empirically based behavior change

techniques [2,9,15,16]. With hindsight, we might have been wiser to have chosen a more focused audience (such as people using heroin or people in early recovery) and then developed and disseminated the app in response to their particular wishes. Likewise, we might have included additional evidence-based behavior change techniques such as goal setting, action planning, and social support [32-34]. These potential changes notwithstanding, we would still have retained our co-design process, given that even very powerful behavior change techniques are undermined if they are delivered in a way that people do not understand, do not trust, or deem boring or unacceptable.

Given the importance of personal and trusted recommendation in relation to uptake, we conclude that an effective dissemination strategy requires a network of supportive partners, that is, credible organizations and individuals who will champion the app by telling others that it exists, linking it to their own websites and informational materials, and proactively disseminating it by means of social media. In addition, it would be helpful if these partners were able to offer guidance and support to people who might use the app to ensure that they know how to download it, understand all the functions, and are able to capitalize on what is being offered [11]. More generally, our findings remind us that apps need to be maintained and regularly updated in response to user feedback. This requires resources (money, time, and expertise) alongside a business model for sustainability [1]. In addition, consideration needs to be given to *competitor apps*. Some of our participants stated that there were other apps for sleep; therefore, they were not interested in using SUSS. Over time, we will likely see more free recovery apps published, some of which will probably have additional capacities (for example, GPS, motion sensors, biophysical monitoring, 24-hour professional support, or linkages to primary care–based treatment) [1,7,15]. Although we appreciate that some potential app users may feel overwhelmed or confused by having too much choice, this is unlikely to be a problem with respect to recovery apps aimed at people experiencing alcohol or other drug problems where options are currently very limited. We therefore reject the view that *competitor apps* are a problem. Instead, we feel that having more well-designed apps that seek to support people in overcoming problems with their substance use indicates that this is a viable space for technological innovation and the availability of a pool of apps should provide welcome choice for an often-underserved population.

Turning finally to how our findings have started to shape and influence our own work, we have recently begun to develop and support a community of SURE Recovery users. To this end, we have recruited a small group of people (SURE Recovery champions) who have lived experience of substance use, have good information technology skills, and work or volunteer in addiction services in different areas of the United Kingdom. These individuals have been provided with a tablet computer, training in how to use and explain SURE Recovery to others, a small honorarium, and out-of-pocket expenses to enable them to travel to services and demonstrate the app. The champions also meet monthly on the web with members of the core app development team to share knowledge and understanding. This initiative has been established to capitalize on our finding that people are more interested in the app when they hear about it from a trusted source. In addition, by connecting the tablet to the free Wi-Fi within services, the SURE Recovery champions increase the app's accessibility to people who may not have hardware or their own mobile data plans.

Beyond this, the dynamic *hot topic of the month* feature introduced in March 2021 has enabled us to begin to balance out the content of SURE Recovery by adding more questions and resources relating to alcohol and wider topics that are not opioid specific (such as mental health, diet and eating, peer support, mindfulness, and stigma). We have also used the hot topic feature to invite app users to provide us with *words of wisdom* to pass on to other people in recovery, and some of these reflections have now been included in the banner of the app. This responds to requests to include inspirational quotations, while also helping to make the app more participative and increasing the feeling of community among users. To supplement this, we have amplified our social media presence with a Facebook page, YouTube channel, Twitter handle, and Instagram account. Going forward, we will also be meeting again with our app developer to discuss the inclusion of push notifications and reminders, as well as brainstorming other ideas for updating the app based on our research findings.

Limitations

Our study and analyses inevitably include limitations. The research was conducted by members of the team who developed SURE Recovery. Our findings may consequently suffer from social desirability bias [35] because the people who were interviewed might not have been as critical as they would have been if the research team had been wholly independent of the app. Equally, we only interviewed 20 people. Although we were careful in selecting participants who reported different levels of engagement with SURE Recovery, we recognize that those interviewed were not necessarily representative of people downloading or using our app. Furthermore, people who download and use SURE Recovery are not representative of all people using substances. This is clear from the fact that our interview participants were all aged 25-63 years and identified as White British. That our findings do not capture the views and experiences of people of color is a particular shortcoming within a field that has historically underrepresented populations identifying as non-White. We hope that others developing and evaluating addiction-related apps will learn from this limitation and will consider collecting and analyzing ethnicity data to better understand if and why some populations may not engage and to help ensure that future apps are clearly relevant to a range of ethnic and racial groups. Because of these limitations, we cannot claim that our findings reflect the views and experiences of people with different ethnicities and demographic characteristics. Nonetheless, we are encouraged by the fact that key patterns and themes identified in our data are found in other research. This provides a degree of reassurance that our findings, although not empirically generalizable, are likely to be relevant and *transfer* to other related apps and settings [27,28].

Conclusions

To conclude, we return to the question of whether there is a genuine future for mHealth apps aimed at people in recovery from alcohol and other drug problems and, if so, what can be done to promote uptake and engagement. Our findings are cautiously positive but show that additional effort is needed. Although first impressions, trusted recommendations, personal relevance, and perceived benefits will all play a role, addiction recovery app uptake and engagement continue to be undermined by broader structural issues of digital exclusion and marginalization [36,37]. Until there is wider access to devices and mobile data and better universal information technology literacy, the potential of any digital intervention is not likely to be achieved [14]. For the foreseeable future, we will therefore need addiction service providers to support mHealth interventions by providing access to devices, onsite Wi-Fi, and training and support in using digital technology. Meanwhile, people who do use apps will not always be expecting them to reduce their substance use or increase their days sober. They may also download and engage with an app because it looks interesting, makes them feel better, lifts their mood, helps them to feel connected with others, or is fun. Accordingly, we should not assess the success of addiction recovery apps based only on objective measures of changes in substance use. As an app can only ever be a cog in a wider ecosystem of support and treatment, we also need to judge its impact through more subjective indicators of health and well-being that may be transient and difficult to quantify but important nonetheless.

Acknowledgments

The authors would like to thank all study participants for sharing their experiences and all members of the wider SURE Recovery project team, particularly Mel Getty, Paul Lennon, and Professor Sir John Strang, for their ongoing support and involvement.

Authors' Contributions

JN was responsible for conceptualization, methodology, data curation, formal analysis, supervision, and writing the original draft. AMB was responsible for methodology, investigation, data curation, formal analysis, and writing the original draft.

Conflicts of Interest

In the last 3 years, JN has received, through her university, research funding from Mundipharma Research Ltd and Camurus AB (for unrelated research) and an honorarium from Indivior (for an unrelated conference presentation).

References

1. Ford JH, Alagoz E, Dinauer S, Johnson KA, Pe-Romashko K, Gustafson DH. Successful organizational strategies to sustain use of A-CHESS: a mobile intervention for individuals with alcohol use disorders. J Med Internet Res 2015 Aug 18;17(8):e201 [FREE Full text] [doi: 10.2196/jmir.3965] [Medline: 26286257]
2. Hoeppner BB, Schick MR, Kelly LM, Hoeppner SS, Bergman B, Kelly JF. There is an app for that - Or is there? A content analysis of publicly available smartphone apps for managing alcohol use. J Subst Abuse Treat 2017 Nov;82:67-73. [doi: 10.1016/j.jsat.2017.09.006] [Medline: 29021117]
3. Copeland J. A qualitative study of barriers to formal treatment among women who self-managed change in addictive behaviours. J Subst Abuse Treat 1997;14(2):183-190. [doi: 10.1016/s0740-5472(96)00108-0] [Medline: 9258863]
4. Cunningham JA, Sobell LC, Sobell MB, Agrawal S, Toneatto T. Barriers to treatment: why alcohol and drug abusers delay or never seek treatment. Addictive Behav 1993 May;18(3):347-353. [doi: 10.1016/0306-4603(93)90036-9]
5. Grant BF. Barriers to alcoholism treatment: reasons for not seeking treatment in a general population sample. J Stud Alcohol 1997 Jul;58(4):365-371. [doi: 10.15288/jsa.1997.58.365] [Medline: 9203117]
6. Marsch LA. Leveraging technology to enhance addiction treatment and recovery. J Addict Dis 2012;31(3):313-318 [FREE Full text] [doi: 10.1080/10550887.2012.694606] [Medline: 22873192]
7. Meredith SE, Alessi SM, Petry NM. Smartphone applications to reduce alcohol consumption and help patients with alcohol use disorder: a state-of-the-art review. Adv Health Care Technol 2015;1:47-54 [FREE Full text] [doi: 10.2147/AHCT.S65791] [Medline: 27478863]
8. Savic M, Best D, Rodda S, Lubman DI. Exploring the focus and experiences of smartphone applications for addiction recovery. J Addict Dis 2013;32(3):310-319. [doi: 10.1080/10550887.2013.824331] [Medline: 24074196]
9. Ramo DE, Popova L, Grana R, Zhao S, Chavez K. Cannabis mobile apps: a content analysis. JMIR Mhealth Uhealth 2015 Aug 12;3(3):e81 [FREE Full text] [doi: 10.2196/mhealth.4405] [Medline: 26268634]
10. Chapman C, Champion KE, Birrell L, Deen H, Brierley M, Stapinski LA, et al. Smartphone apps about crystal methamphetamine ("Ice"): systematic search in app stores and assessment of composition and quality. JMIR Mhealth Uhealth 2018 Nov 21;6(11):e10442 [FREE Full text] [doi: 10.2196/10442] [Medline: 30463834]
11. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 May 29;22(5):e17572 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]

12. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
13. Gray R, Kelly P, Beck A, Baker A, Deane F, Neale J, et al. A qualitative exploration of SMART recovery meetings in Australia and the role of a digital platform to support routine outcome monitoring. Addict Behav 2020 Feb;101:106144. [doi: 10.1016/j.addbeh.2019.106144] [Medline: 31648139]
14. Milward J, Day E, Wadsworth E, Strang J, Lynskey M. Mobile phone ownership, usage and readiness to use by patients in drug treatment. Drug Alcohol Depend 2015 Jan 01;146:111-115. [doi: 10.1016/j.drugalcdep.2014.11.001] [Medline: 25468818]
15. Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L. Smartphone apps targeting alcohol and illicit substance use: systematic search in in commercial app stores and critical content analysis. JMIR Mhealth Uhealth 2019 Apr 22;7(4):e11831 [FREE Full text] [doi: 10.2196/11831] [Medline: 31008713]
16. Penzenstadler L, Chatton A, Van Singer M, Khazaal Y. Quality of smartphone apps related to alcohol use disorder. Eur Addict Res 2016;22(6):329-338. [doi: 10.1159/000449097] [Medline: 27598779]
17. King's College London. URL: https://www.kcl.ac.uk/research/sure-recovery-app [accessed 2022-02-18]
18. Neale J, Vitoratou S, Finch E, Lennon P, Mitcheson L, Panebianco D, et al. Development and validation of 'SURE': a patient reported outcome measure (PROM) for recovery from drug and alcohol dependence. Drug Alcohol Depend 2016 Aug 01;165:159-167 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.06.006] [Medline: 27344196]
19. Neale J, Vitoratou S, Lennon P, Meadows R, Nettleton S, Panebianco D, et al. Development and early validation of a patient-reported outcome measure to assess sleep amongst people experiencing problems with alcohol or other drugs. Sleep 2018 Apr 01;41(4):zsy013 [FREE Full text] [doi: 10.1093/sleep/zsy013] [Medline: 29329423]
20. Islind AS, Lindroth T, Lundin J, Steineck G. Co-designing a digital platform with boundary objects: bringing together heterogeneous users in healthcare. Health Technol 2019 May 20;9(4):425-438. [doi: 10.1007/s12553-019-00332-5]
21. Kensing F, Greenbaum J. Heritage: having a say. In: Routledge International Handbook of Participatory Design. Oxfordshire, England, UK: Routledge; 2012.
22. Visser FS, Stappers PJ, van der Lugt R, Sanders EB. Contextmapping: experiences from practice. CoDesign 2005 Apr;1(2):119-149. [doi: 10.1080/15710880500135987]
23. Trischler J, Pervan SJ, Kelly SJ, Scott DR. The value of codesign. J Service Res 2017 Jul 10;21(1):75-100. [doi: 10.1177/1094670517714060]
24. 11 lessons: a study of the design process. Design Council. URL: https://www.designcouncil.org.uk/resources/report/11-lessons-managing-design-global-brands [accessed 2022-02-18]
25. Neale J, Pickering L, Nettleton S. The everyday lives of recovering heroin users. Royal Society of Arts. URL: https://www.thersa.org/reports/everyday-lives [accessed 2022-02-18]
26. Glaser B, Strauss A. The Discovery of Grounded Theory Strategies for Qualitative Research. Piscataway, New Jersey: Aldine Transaction; 1999.
27. Neale J. Iterative categorization (IC): a systematic technique for analysing qualitative data. Addiction 2016 Jun;111(6):1096-1106 [FREE Full text] [doi: 10.1111/add.13314] [Medline: 26806155]
28. Neale J. Iterative categorisation (IC) (part 2): interpreting qualitative data. Addiction 2021 Mar;116(3):668-676. [doi: 10.1111/add.15259] [Medline: 32926762]
29. Innovation in qualitative and mixed methods data analysis. MAXQDA. URL: https://www.maxqda.com/about [accessed 2022-02-18]
30. Beck AK, Kelly PJ, Deane FP, Baker AL, Hides L, Manning V, et al. Developing a mHealth routine outcome monitoring and feedback app ("SMART track") to support self-management of addictive behaviours. Front Psychiatry 2021;12:677637 [FREE Full text] [doi: 10.3389/fpsyt.2021.677637] [Medline: 34220583]
31. 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]
32. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
33. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
34. Michie S, Johnston M. Optimising the value of the evidence generated in implementation science: the use of ontologies to address the challenges. Implement Sci 2017 Nov 14;12(1):131 [FREE Full text] [doi: 10.1186/s13012-017-0660-2] [Medline: 29137660]
35. Grimm P. Social desirability bias. Wiley Int Encyclopedia Market 2010 [FREE Full text]
36. Neale J, Stevenson C. Homeless drug users and information technology: a qualitative study with potential implications for recovery from drug dependence. Subst Use Misuse 2014 Sep;49(11):1465-1472. [doi: 10.3109/10826084.2014.912231] [Medline: 24799073]

37. Longley PA, Singleton AD. Linking social deprivation and digital exclusion in England. Urban Stud 2009 May 06;46(7):1275-1298. [doi: 10.1177/0042098009104566]

Abbreviations

mHealth: mobile health
SURE: Substance Use Recovery Evaluator
SUSS: Substance Use Sleep Scale

Edited by A Kushniruk; submitted 23.08.21; peer-reviewed by B Hoeppner, D Szinay; comments to author 07.11.21; revised version received 25.11.21; accepted 19.12.21; published 08.03.22.

Please cite as:
Neale J, Bowen AM
Lessons for Uptake and Engagement of a Smartphone App (SURE Recovery) for People in Recovery From Alcohol and Other Drug Problems: Interview Study of App Users
JMIR Hum Factors 2022;9(1):e33038
URL: https://humanfactors.jmir.org/2022/1/e33038
doi:10.2196/33038
PMID:35258474

Original Paper

Suitability of the Unified Theory of Acceptance and Use of Technology 2 Model for Predicting mHealth Acceptance Using Diabetes as an Example: Qualitative Methods Triangulation Study

Patrik Schretzlmaier[1], MEng; Achim Hecker[2,3], PhD; Elske Ammenwerth[1], PhD

[1]Institute of Medical Informatics, UMIT – Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria

[2]Institute for Management and Economics in Healthcare, UMIT – Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria

[3]DBU Digital Business University of Applied Sciences, Berlin, Germany

Corresponding Author:
Patrik Schretzlmaier, MEng
Institute of Medical Informatics
UMIT – Private University for Health Sciences, Medical Informatics and Technology
Eduard-Wallnoefer-Zentrum 1
Hall in Tirol, 6060
Austria
Phone: 49 17657875980
Email: patrik.schretzlmaier@edu.umit-tirol.at

Abstract

Background: In recent years, the use of mobile health (mHealth) apps to manage chronic diseases has increased significantly. Although mHealth apps have many benefits, their acceptance is still low in certain areas and groups. Most mHealth acceptance studies are based on technology acceptance models. In particular, the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) model was developed to predict technology acceptance in a consumer context. However, to date, only a few studies have used the UTAUT2 model to predict mHealth acceptance and confirm its suitability for the health sector. Thus, it is unclear whether the UTAUT2 model is suitable for predicting mHealth acceptance and whether essential variables for a health-related context are missing.

Objective: This study aims to validate the suitability of UTAUT2 for predicting mHealth acceptance.

Methods: In this study, diabetes was used as an example as mHealth apps are a significant element of diabetes self-management. In addition, diabetes is one of the most common chronic diseases affecting young and older people worldwide. An explorative literature review and guided interviews with 11 mHealth or technology acceptance experts and 8 mHealth users in Austria and Germany were triangulated to identify all relevant constructs for predicting mHealth acceptance. The interview participants were recruited by purposive sampling until theoretical saturation was reached. Data were analyzed using structured content analysis based on inductive and deductive approaches.

Results: This study was able to confirm the relevance of all exogenous UTAUT2 constructs. However, it revealed two additional constructs that may also need to be considered to better predict mHealth acceptance: *trust* and *perceived disease threat*.

Conclusions: This study showed that the UTAUT2 model is suitable for predicting mHealth acceptance. However, the model should be extended to include 2 additional constructs for use in the mHealth context.

(JMIR Hum Factors 2022;9(1):e34918) doi:10.2196/34918

KEYWORDS

mHealth; mobile health; mobile apps; diabetes mellitus; technology acceptance; UTAUT2; mobile phone

Introduction

Background

Mobile health (mHealth) apps are essential for effective self-management of chronic diseases such as diabetes [1,2]. In this context, mHealth describes mHealth technologies such as diabetes apps [3] and continuous glucose monitoring (CGM) systems [4] to support diabetes self-management and patient health [5-11]. The use of mHealth apps for diabetes self-management leads to more frequent monitoring of blood glucose levels and lower long-term glucose levels [12]. However, many patients do not use mHealth apps as they do not see the necessity or are satisfied with their current management [13]. Despite the potential and relevance of mHealth apps in chronic disease management, they are still used insufficiently [14].

An important aspect that determines the use of mHealth apps is their acceptance [15,16]: “User acceptance can be defined as the demonstrable willingness within a user group to employ information technology for the tasks it is designed to support" [17]. However, the acceptance of mHealth apps is still low in certain areas and groups [5,18-23]. For example, for type 2 diabetes, the acceptance of mHealth apps is low [12,24].

User acceptance often determines the success or failure of technical apps [25]. For predicting the acceptance of mHealth users, technology acceptance models are used [16,25]. These models are essential as they combine various theories from psychology and sociology to explain and predict technology acceptance and use [26].

We used diabetes as an example to investigate this issue as mHealth apps are a significant element of diabetes self-management [2,4,7].

In addition, diabetes is one of the most common chronic diseases, affecting approximately 463 million people worldwide between the ages of 20 and 79 years in 2019 [27]. Most patients (approximately 90%) have type 2 diabetes [27], where effective self-management can have a significant impact on improving patient health [5,6].

Many mHealth apps such as smartphone apps, blood glucose sensors (CGM), and others are used by patients with type 1 and type 2 diabetes in their self-management.

Therefore, we wanted to investigate the following research question in the field of mHealth self-management in diabetes: is the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) model suitable for predicting mHealth acceptance using diabetes as an example?

If the UTAUT2 model was better adapted to the needs of mHealth acceptance, the reasons for use or rejection of mHealth apps could thus be better predicted and more easily taken into account in new developments. This would help increase the use of mHealth self-management apps among people who are chronically ill, thereby improving their health.

Theoretical Background

In health informatics, the Technology Acceptance Model (TAM), UTAUT, and UTAUT2 have proven to be suitable models for acceptance research [28-30]. These models consider constructs that influence the acceptance of technology to predict its use [28].

The TAM was developed in the late 1980s and provided the basis for further technology acceptance models [16,25]. It focuses on understanding why users accept or reject information technology (IT) systems and how their design influences acceptance [25]. The TAM hypothesizes that *perceived usefulness* and *perceived ease of use* are essential for the *attitude toward using*, which is a dominant factor of *behavioral intention to use* and can be interpreted as technology acceptance [25,28,31].

In 2003, the UTAUT model was published to present a unified model that synthesizes the diversity of acceptance models [16]. The basis of the UTAUT model is the analysis and comparison of 8 technology acceptance models (eg, Theory of Planned Behavior, TAM, and Innovation Diffusion Theory [16]). The UTAUT model aims to evaluate the likelihood of success of new technologies and understand the critical acceptance factors to proactively define measures to ensure that systems are accepted and used [16]. It uses the four central constructs of performance expectancy, effort expectancy, social influence, and facilitating conditions, moderated by gender, age, experience, and voluntariness of use, as direct determinants of behavioral intention and use behavior [16].

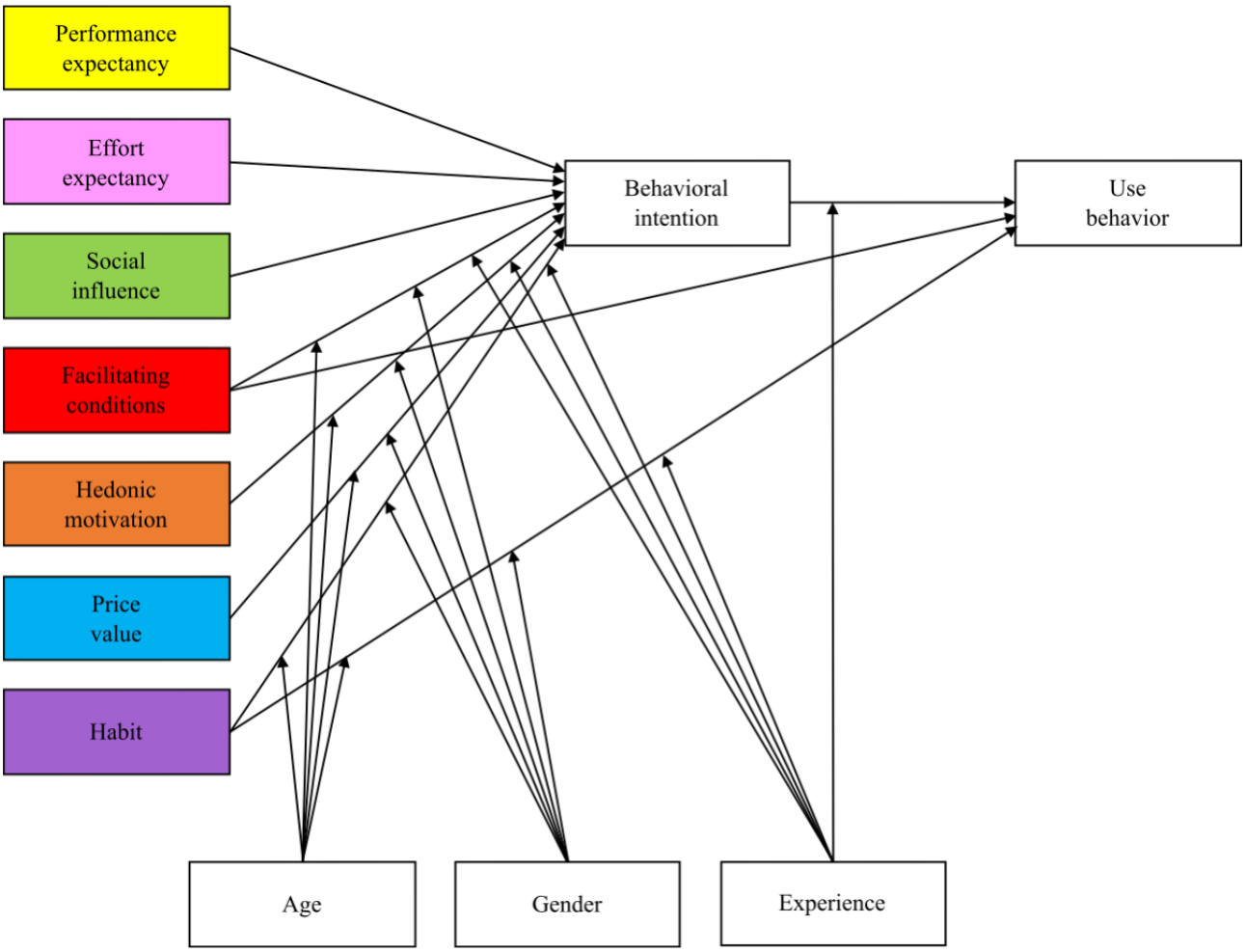
To date, mHealth acceptance studies have mainly used the TAM [32-35] and UTAUT [19,36,37] model or combinations of both [38,39]. Although the TAM was developed to predict the acceptance of IT systems [25], the UTAUT model focused on behavioral intention and technology use in organizational contexts [16].

In contrast, the focus of the UTAUT2 model, which was developed as an extension of the UTAUT model, is to predict technology acceptance in consumer use contexts [26]. Therefore, additional constructs such as hedonic motivation, price value, and habit were added [26].

Figure 1 shows the UTAUT2 model developed by Venkatesh et al [26], with its exogenous constructs (colored boxes) of performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit. It also shows the relationships between these exogenous constructs and the endogenous constructs of behavioral intention and use behavior [26]. Some of these relationships are moderated by age, gender, and experience [26].

As this study focuses on the suitability of the UTAUT2 model for predicting mHealth acceptance, Textbox 1 shows the definitions of the exogenous UTAUT2 constructs only, adapted from the study by Venkatesh et al [26].

Figure 1. The Unified Theory of Acceptance and Use of Technology 2 model, adapted from a study by Venkatesh et al [26].



Textbox 1. Exogenous Unified Theory of Acceptance and Use of Technology 2 constructs adapted from a study by Venkatesh et al [26].

Performance expectancy

- "Degree to which using a technology will provide benefits to consumers in performing certain activities" [26]

Effort expectancy

- "Degree of ease associated with consumers' use of technology" [26]

Social influence

- "Extent to which consumers perceive that important others (e.g. family and friends) believe they should use a particular technology" [26]

Facilitating conditions

- "Refer to consumers' perceptions of the resources and support available to perform a behavior" [26]

Hedonic motivation

- "The fun or pleasure derived from using a technology" [26]

Price value

- "Consumers' cognitive tradeoff between the perceived benefits of the applications and the monetary cost for using them" [26]

Habit

- "The extent to which people tend to perform behaviors automatically because of learning" [26]

Therefore, the UTAUT2 model seems appropriate specifically for mHealth technologies as it focuses on individuals and their needs [26,40]. This is visible, for example, in the construct of hedonic motivation, which has been described in some scientific articles as particularly important for consumers of a product or technology [26,41,42].

However, to date, only a few mHealth acceptance studies have used the UTAUT2 model, and out of the studies using it, some showed that primarily health-related factors such as health conditions, health consciousness, and health concerns are missing from the technology acceptance model [41,43,44]. These are particularly relevant for patients with chronic diseases who are using mHealth apps. In this context, mHealth acceptance may depend not only on fun or habit but also on the perceived threat of disease and perceived data security [18,19]. However, these aspects are not covered in the UTAUT2 model.

Methods

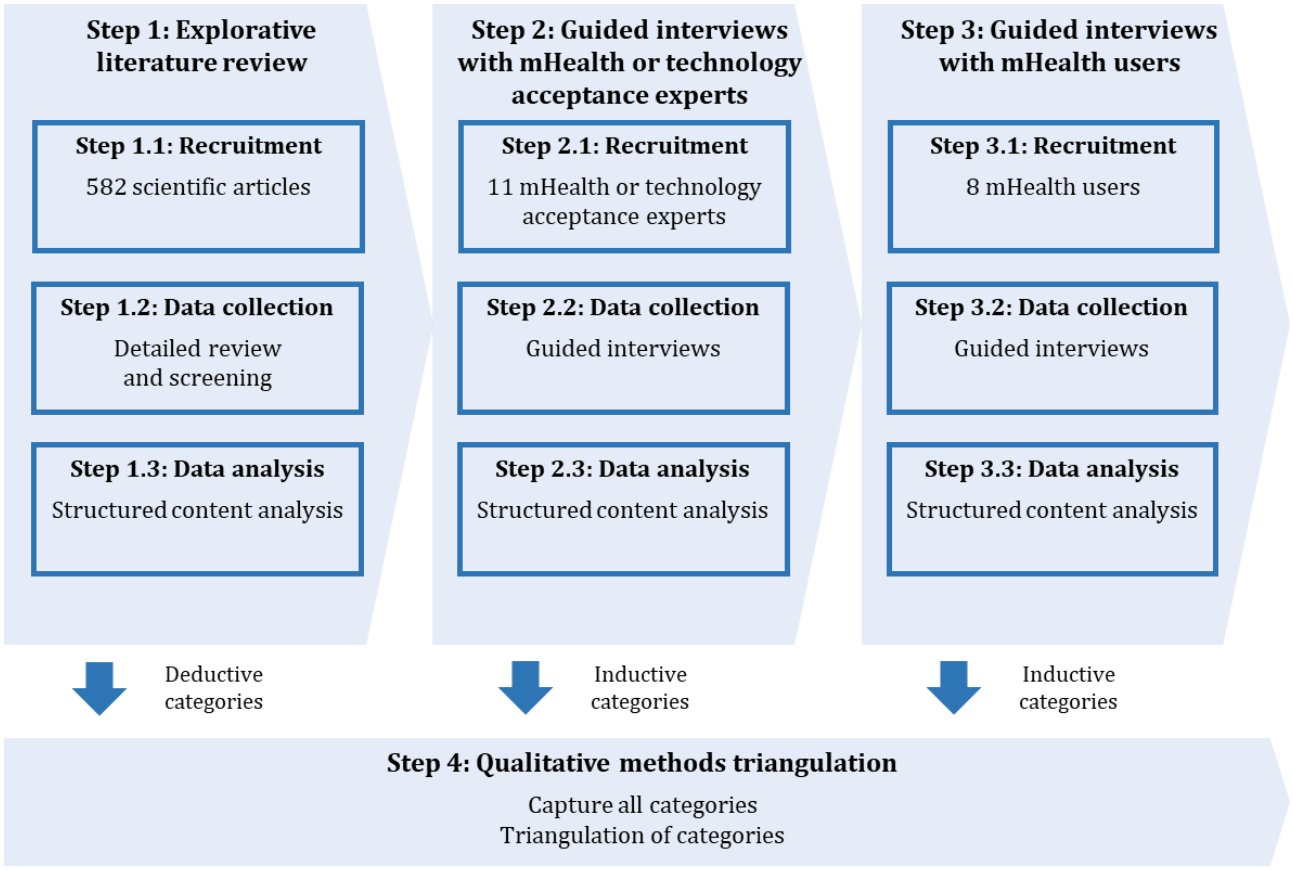
We followed the 32-item COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [45].

Design

We used a qualitative research design and triangulated an explorative literature review with guided interviews. The objective was to identify the main categories of mHealth acceptance in the field of diabetes self-management.

The research design used, as shown in Figure 2, comprises 4 main steps (step 1 to step 4) built on each other. In the first step (step 1), we identified relevant categories from the explorative literature review for the initial category system. In the second step (step 2), we conducted guided interviews with mHealth or technology acceptance experts, followed by the third step (step 3), where we conducted guided interviews with mHealth users. Guided interviews and literature review served to assess the existing exogenous UTAUT2 constructs in a health-related context and identify possible additional categories. In the last step (step 4) of the research process, we used qualitative methods triangulation to capture and compare all identified categories from the previous research steps (step 1 to step 3) and finally confirmed or rejected them.

Figure 2. Research design. mHealth: mobile health.



Ethics Approval

The study was approved by the research committee for scientific ethical questions of the UMIT Private University for Health Sciences, Medical Informatics and Technology (reference number RCSEQ 2805/20).

Explorative Literature Review (Step 1)

Recruitment (Step 1.1)

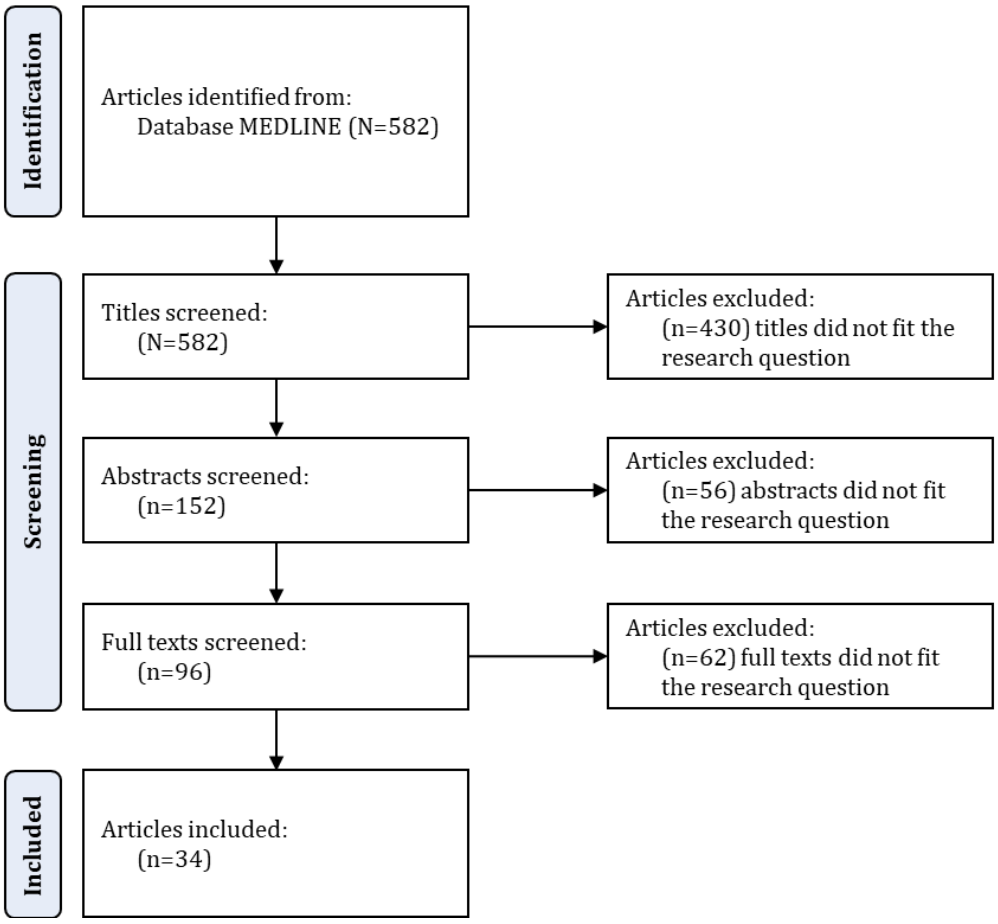
Between March and November 2020, we conducted an explorative literature review in the MEDLINE database following systematic criteria. We used the keywords *diabetes* and *diabetes mellitus* for the concept of diabetes. For the concept of mHealth apps, we used the keywords *mobile health apps*,

mobile health applications, *mobile health units*, and *mobile apps*. For the concept of technology acceptance, we used the keywords *acceptance*, *UTAUT*, and *UTAUT2*. In total, we identified 582 scientific articles using different search queries.

Data Collection (Step 1.2)

The explorative literature review aimed to identify relevant scientific articles from the mHealth and technology acceptance field to develop the initial category system based on the UTAUT2 model and additional categories using diabetes as an example. On the basis of the identified scientific articles, we conducted a screening process, which is described in Figure 3. In the screening process, we first checked the titles and then the abstracts of all scientific articles and compared them with the research question. In these 2 steps, of the 582 scientific articles, we filtered out 486 (83.5%) scientific articles that did not meet the inclusion criteria, and for the remaining 96 (16.5%) scientific articles, we conducted a full-text analysis and compared the content of the methods, results, and conclusions sections with the research question. Approximately 5.8% (34/582) of scientific articles met the inclusion criteria.

Figure 3. Explorative literature review—the screening process.



Data Analysis (Step 1.3)

We conducted data analysis sequentially for each research step (step 1 to step 3). MAXQDA 2020 (release 20.4.0; VERBI GmbH) was used for transcribing and coding the qualitative data. We conducted a structured content analysis using inductive and deductive approaches, following the research question to analyze the qualitative material according to Kuckartz [46]. We developed, used, and continuously updated a codebook containing relevant information (eg, detailed code description and inclusion and exclusion criteria) to ensure the high quality of the coding process [47]. Throughout the complete data analysis process (step 1.3, step 2.3, and step 3.3), we used the method of peer debriefing, in which we critically discussed the collected data and the results derived from that data, as well as the related analysis processes with an experienced research expert [48]. In particular, unclear passages in the qualitative data were reviewed according to the four eyes principle and discussed during meetings with the coauthors to find a shared consensus.

We started the data analysis by coding the scientific articles from the explorative literature review (step 1). In the first deductive step, we defined categories based on exogenous UTAUT2 constructs to develop the initial category system. In the second step, we coded 34 scientific articles based on predefined categories and assigned all the relevant text segments to the corresponding categories. In the third step, we inductively defined and added missing categories to the category system based on the material.

Guided Interviews With mHealth or Technology Acceptance Experts (Step 2)

Recruitment (Step 2.1)

We conducted guided interviews between December 2020 and March 2021 with 11 mHealth or technology acceptance experts (9, 82% men and 2, 18% women) from Germany and Austria. We identified the experts based on their publications and institute websites. Each of the experts held, at minimum, a PhD degree and had worked in the research area of mHealth or technology acceptance for ≥3 years. We used purposive sampling to select the experts from universities in Germany and Austria. After the 11 interviews, theoretical saturation was reached.

Data Collection (Step 2.2)

On the basis of the results of the explorative literature review (step 1) and the research question, we developed individual theory-based interview guides with open-ended questions for the interviews with mHealth or technology acceptance experts (step 2) and mHealth users (step 3). We tested and improved the interview guides before the official interviews. The researcher (PS) who conducted the interviews was trained in qualitative research methods and had a positive interest in mHealth apps. There was no personal relationship between the researcher and the interview candidates. The interviews took place only between the researcher and the interview candidate on the web (web conference) or by telephone. Both the researcher and interview candidate were at home or in their own office during the interview; therefore, no one else was present. Before the interview started, there was a short introduction of the researcher, the research topic, and the data privacy guidelines. All interviews were conducted in German, audio recorded by an external audio recording device, and lasted between 20 and 45 minutes. After we finished the interviews, we offered the participants the opportunity to ask questions, which helped improve the interview guides. We took notes on the interview atmosphere and comments from outside the interview. We did not repeat any interviews, and there were no dropouts.

We started the guided interviews with mHealth or technology acceptance experts (step 2). The interviews aimed to assess the existing exogenous UTAUT2 constructs in a health-related context and identify additional relevant categories. Therefore, in the first part of the interviews, we asked questions about the general factors influencing the acceptance and sustained use of mHealth apps from an expert perspective: *which factors significantly affect the acceptance and long-term use of mHealth self-management apps?* In the second part of the interviews, we focused on the UTAUT2 model, specifically on the essential constructs and constructs that should be added based on the experts' feedback: *which constructs should be supplemented to the UTAUT2 model concerning acceptance investigations of mHealth self-management apps?* (Table 1). For this reason, we adopted an unprompted approach with open-ended questions.

Table 1. Main topics of the guided interviews with mHealth[a] or technology acceptance experts (n=11) and mHealth users (n=8).

mHealth or technology acceptance experts	mHealth users
General	
Factors influencing the acceptance and long-term use of mHealth self-management apps	Factors influencing the (long-term) use of mHealth self-management apps
Advantages and disadvantages associated with the use of mHealth self-management apps	Advantages and disadvantages associated with the use of mHealth self-management apps
Reasons leading to use or nonuse of mHealth self-management apps	Reasons leading to use or nonuse of mHealth self-management apps
Specific	
UTAUT2[b] variables have the most significant influence on the acceptance and use	Expectations, barriers, and emotions related to the use of the mHealth self-management app over time
Variables that should be added to the UTAUT2 model to describe the acceptance of mHealth self-management apps	Relevance of the mHealth self-management app in daily life

[a]mHealth: mobile health.

[b]UTAUT2: Unified Theory of Acceptance and Use of Technology 2.

Data Analysis (Step 2.3)

In contrast to the data analysis of the explorative literature review (step 1.3), the analysis of the guided interviews (step 2.3 and step 3.3) was not conducted at the end of the entire data collection phase but continuously after each interview. This iteration process helped us identify the point of theoretical saturation; that is, the point at which we were no longer able to identify new categories [48].

In the first step, we continued the data analysis by transcribing the guided interviews with mHealth or technology acceptance experts verbatim. We did not return the interview transcripts to the participants. In the second step, we coded each interview based on the differentiated category system containing the deductive and inductive categories, which resulted from the explorative literature review (step 1.3). In the third step, we inductively defined and added missing categories to the category system based on the material until saturation was reached.

Guided Interviews With mHealth Users (Step 3)

Recruitment (Step 3.1)

Between March and May 2021, we conducted guided interviews with 8 mHealth users (5, 63% men and 3, 38% women) from Germany and Austria. The age distribution of the participants ranged from 20 to 75 years. We included patients with type 1 and type 2 diabetes and parents caring for children with type 1 diabetes, as the requirements and needs for mHealth apps are comparable, and the apps do not specifically address only one user group. We only included participants using an mHealth app (diabetes app and CGM system) for at least 3 months.

We identified mHealth users through gatekeepers in organizations such as diabetes associations and diabetes self-help groups, who asked suitable persons to participate in the study. In addition, we published a call for participation in the study on social media. We used purposive sampling to recruit patients of different ages, genders, and socioeconomic backgrounds to ensure a wide diversity (Table 2). After 8 interviews, theoretical saturation was reached.

Table 2. Sociodemographic data of recruited mHealth[a] users.

User	Age (years)	Gender	Education	Residence	Type of diabetes	Duration of mHealth app use
1	75	Female	PhD	Austria	Type 2	4 months
2	33	Female	Vocational qualification	Germany	Type 1	4 years
3	52	Male	PhD	Austria	Type 2	6 months
4	20	Female	Vocational qualification	Germany	Type 1	2 years
5	40	Male	PhD	Austria	Father of type 1 diabetes child	4 years
6	22	Male	Student	Germany	Type 1	3 years
7	23	Male	Student	Austria	Type 1	6 years
8	60	Male	Master's	Austria	Type 2	4 months

[a]mHealth: mobile health.

Data Collection (Step 3.2)

After we completed all the interviews with mHealth or technology acceptance experts (step 2), we continued the interviews with mHealth users (step 3). The interviews also aimed to confirm the existing exogenous UTAUT2 constructs and identify additional relevant categories based on mHealth users' perspectives. The first part of the interviews focused on the used mHealth app and the reasons for choosing and using it: *which features or functions are essential to you so that you use diabetes mHealth self-management apps in the long term?* In the second part of the interviews, we focused on the users' experience with the mHealth app: *when you first started to use diabetes mHealth self-management apps, what expectations did you have?* (Table 1). For this purpose, we used the user experience (UX) curve method [49], which visualized the UX throughout use. We drew the UX curve by sharing the screen with the mHealth users who joined the interviews on the web. This was not possible if the interview was conducted via telephone. In those cases, we only asked questions on UX without visualization using the UX curve method.

Data Analysis (Step 3.3)

To analyze the data from the mHealth user interviews (step 3.3), we performed the same analysis steps as for the analysis of the mHealth or technology acceptance expert interviews (step 2.3). However, for coding the mHealth user interviews, we used the already differentiated category system that included inductive categories from the interviews with mHealth or technology acceptance experts (step 2.3). Approximately 2 weeks after the coding of all material from step 1 to step 3 was completed, we reviewed the final category system and the coded segments to ensure the reliability (intrarater reliability) of the analyzed data [50]. Identical or similar categories were combined.

Qualitative Methods Triangulation (Step 4)

We used qualitative methods triangulation [51,52] to combine the different perspectives from the explorative literature review (step 1) and the guided interviews (step 2 and step 3) to investigate the research question, thereby increasing confidence in the results and their validity [48,52]. For this purpose, we captured and compared all identified categories from the 3 research steps (step 1 to step 3) to determine the relevant categories to answer the research question. We considered categories that we identified in at least two of the three research steps (step 1 to step 3) to be particularly important for extending the UTAUT2 model.

Results

Overview

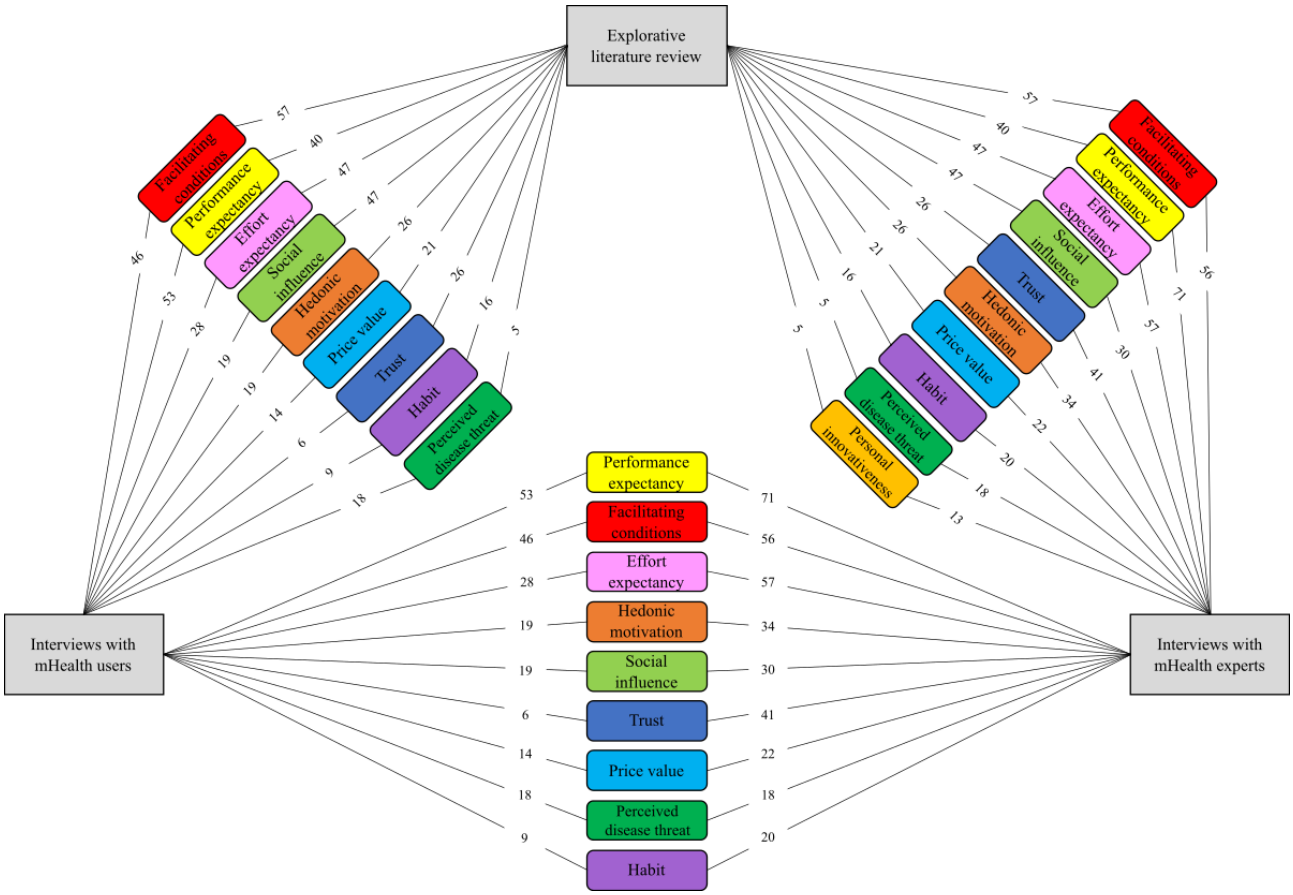
We conducted a qualitative methods triangulation study comprising an explorative literature review (step 1) and guided interviews with 11 mHealth or technology acceptance experts (step 2) and 8 mHealth users (step 3). Using diabetes as an example, we investigated whether the UTAUT2 model is suitable for predicting mHealth acceptance. Thus, we analyzed the material from the explorative literature review (step 1) and the guided interviews (step 2 and step 3) using structured content analysis and then combined the results using qualitative methods triangulation (step 4), as shown in Figure 4.

In our qualitative methods triangulation study, we were able to confirm the relevance of all exogenous UTAUT2 constructs in predicting mHealth acceptance using diabetes as an example.

In addition, we were able to identify another three categories that are not part of the UTAUT2 model: trust, perceived disease threat, and personal innovativeness.

Interview quotes translated verbatim are shown in the following sections to support our results.

Figure 4. Summary of combined categories (colored boxes) identified from explorative literature review and guided interviews (gray boxes). The figures between gray and colored boxes indicate the number of coded segments assigned to each category. Categories are arranged in decreasing order according to the sum of coded segments from both sources. mHealth: mobile health.



Confirmation of Exogenous UTAUT2 Constructs

Overview

According to most (9/11, 82%) of the mHealth or technology acceptance experts interviewed, the UTAUT2 model is suitable for acceptance studies, especially in areas where motivation and continuous and voluntary use are essential. For example, this applies to mHealth self-management in diabetes:

> *...in my opinion, it can be used well for all things that are based on voluntariness in the broadest sense...that means especially with health apps that are not compulsory...even if the doctor prescribes me the app free of charge that does not mean that I will use it for the next weeks and months...* [Expert 3, male]

Therefore, as a first essential part of our study, we wanted to confirm that the exogenous UTAUT2 constructs are suitable for predicting mHealth acceptance using diabetes as an example. We summarized the main results of the structured content analysis focusing on the exogenous UTAUT2 constructs to present their relevance in the following sections.

Facilitating Conditions

We were able to confirm the importance of facilitating conditions in diabetes mHealth self-management. In the explorative literature review, we identified several scientific articles that pointed out the relevance of facilitating conditions in mHealth self-management [5,18,21,53-58]. In particular, the authors highlighted technical support, support from the mHealth app itself, and health care professionals as the essential aspects of facilitating conditions for mHealth apps. We were also able to identify those aspects in the guided interviews, as shown in the extracts in the following section.

According to all (8/8, 100%) of the mHealth users surveyed, good technical support, especially if there are any problems, and support from their medical physician are essential for accepting mHealth apps:

> *...it is crucial to me in any case, especially with technical problems when the sensor, the mechanism is broken, that you are told how to place it or that you can easily contact the support...In any case, it is vital to me that the doctors can get a good picture and simply that the disease is kept under control.* [User 7, male, type 1 diabetes]

Some (4/11, 36%) of the mHealth or technology acceptance experts confirmed that support from medical physicians is an essential factor for the long-term use of mHealth self-management apps:

> *...many people also want to have some kind of connection with their doctor. The app is used or recommended by the doctor, or the doctor can be contacted if there are any questions. So this is not just pure self-help, but there is some connection with the healthcare system or with health service...* [Expert 9, female]

Of course, the support from the mHealth app itself is also essential for its use. Some (3/8, 38%) of the mHealth users consider a decent help function (eg, frequently asked questions and video tutorials) that facilitates the use of the mHealth self-management app to be an essential feature:

> *...with pictures and text, there are even videos...every time you put a new sensor...you get an explanation how to do it...* [User 4, female, type 1 diabetes]

Performance Expectancy

We were also able to confirm the importance of performance expectancy in diabetes mHealth self-management. On the basis of the explorative literature review, we identified several scientific articles that showed the importance of performance expectancy in the context of mHealth self-management [18,21,22,55,59-61]. In particular, the authors highlighted the benefits gained for disease management as a relevant aspect of mHealth apps. We were also able to identify this aspect of performance expectancy in the guided interviews, as shown in the following sections.

Most (7/8, 88%) of the mHealth (CGM system) users started using the technology because of clear expectations that it would improve their lives. The mHealth app makes daily management of the disease easier and gives users back a piece of everyday life as it provides an accessible overview of the relevant blood glucose values, which has a positive effect on acceptance:

> *...we switched to it because it is simply a completely different dimension in diabetes management. You can not compare that with regular blood measurements...with this technology, diabetes is just much easier to handle. You can go about your daily life...* [User 5, male, father of type 1 diabetes child]

Compared with traditional blood glucose monitoring, half (4/8, 50%) of the surveyed mHealth (CGM system) users considered setting alarms and reminders in urgent situations to be one of the most crucial functions of the mHealth app, thus increasing its acceptance:

> *...definitely the alarms. I have an Apple Watch that is also compatible with the system. It gives me messages...when I am hypoglycemic when there are any disorders. What I also find to be a significant advantage is sharing the app with other people. So my partner also has it on his cell phone and sees or gets messages when I am hypoglycemic and can no longer react...* [User 4, female, type 1 diabetes]

This also reflects the statements of all (11/11, 100%) of the mHealth or technology acceptance experts. The essential factors for long-term use of mHealth self-management apps are the perceived benefits and advantages that must be visible to the patients, especially the freedom gained and the flexibility in self-management of the disease:

> *...so people have to see that they have a benefit somehow. So over a longer period, they also use it consistently in everyday life...that is the case with chronic diseases, where it is a long-term problem, and you have to make people aware of the benefits of this app for the long term. It must be useful...* [Expert 9, female]

Effort Expectancy

We were also able to confirm the importance of effort expectancy in diabetes mHealth self-management. In the explorative literature review, we were able to identify several scientific articles that showed the importance of effort expectancy in the context of mHealth self-management [5,18,21,40,53,55,57,59,61,62]. In particular, the authors highlighted convenience, simplicity, and usability as relevant aspects associated with the acceptance and use of mHealth apps. As shown in the extracts given in the following section, we were also able to identify those aspects of effort expectancy in the guided interviews.

Usability and simplicity of use without physical impairment were deemed to be essential criteria for long-term use of diabetes mHealth self-management apps by all (8/8, 100%) of the mHealth users. A relevant aspect that half (4/8, 50%) of the surveyed mHealth (CGM system) users highlighted is that drawing blood is not necessary for the glucose measurements, which is a great relief in everyday life and improves acceptance:

> *...it is just a lot easier than when you have to go to the break room at work all the time and prick your finger...take your cell phone, hold it up to the sensor, and it shows you the sugar right away...it is just a significant relief, and you have everything in there...you do not have to keep a diary anymore. You have everything in the app. Everything is there...* [User 2, female, type 1 diabetes]

In all (11/11, 100%) of the mHealth or technology acceptance experts' point of view, the decisive factors for the long-term use of mHealth apps are their user-friendliness and the fact that they require less effort, are easy to use, and produce better outcomes than conventional solutions:

> *...especially the usability plays a decisive role...* [Expert 6, male]

Social Influence

We were also able to confirm the importance of social influence in diabetes mHealth self-management. On the basis of the explorative literature review, we identified several scientific articles that showed the importance of social influence in the context of mHealth self-management [5,18,20-22,36,40,53,57,59,63]. In particular, the authors highlighted the importance of recommendations from physicians, medical

professionals, family members, and friends for the use of mHealth apps. We were also able to identify those aspects in the guided interviews, as shown in the extracts in the following sections.

On the basis of feedback from all (8/8, 100%) of the mHealth users, the primary influence to use an mHealth app for disease management is driven by health care providers such as diabetologists, diabetes outpatient clinics, and physicians. If the personal environment is generally very positive about the mHealth app, acceptance is encouraged:

> *...I was only really made aware of this by my diabetologist. So through her, I got to know that, before I did not know that either...* [User 2, female, type 1 diabetes]

In addition, many (8/11, 73%) of the mHealth or technology acceptance experts see significant influence from health care providers for the first and long-term use of mHealth apps. Most (9/11, 82%) of the mHealth or technology acceptance experts also see some influence from the media and closer personal environment, which positively influences acceptance:

> *...so the recommendation by a doctor, by friends, relatives or other persons involved is fundamental...* [Expert 9, female]

Hedonic Motivation

We were also able to confirm the importance of hedonic motivation in diabetes mHealth self-management. On the basis of the explorative literature review, we were able to identify several scientific articles that showed the importance of hedonic motivation in the context of mHealth self-management [18,22,56,59,62]. In particular, the authors emphasized the importance of emotional support for adherence, motivation through goal setting, and playful elements (ie, gamification) for the sustained use of mHealth apps. We were also able to identify those aspects in the guided interviews, as shown in the extracts given in the following section.

All (8/8, 100%) of the mHealth users reported positive emotions, such as the joy of having an app that helps them manage their disease. Some (5/8, 63%) of the mHealth users associate the use of the app with fun, which leads them to check blood glucose much more frequently, for example, which contributed to increasing the acceptance of the mHealth app:

> *...first of all, joy, because it is a significant relief...* [User 2, female, type 1 diabetes]

> *...you have something new,...you want to use it all the time although it is a medical application...I measured blood sugar fifty times a day...just to see how cool it is...* [User 6, male, type 1 diabetes]

In addition, most (9/11, 82%) of the mHealth or technology acceptance experts consider fun, such as through gamification aspects and positive feedback during use, to be vital motivating factors for ensuring that mHealth apps are used for the long term:

> *...hedonic motivation plays a role—of course, it is a decisive factor in whether you use it or not...I also enjoy it...I find gamification exciting, i.e., increasing motivation through such playful elements...* [Expert 5, male]

Price Value

We were also able to confirm the importance of price value in diabetes mHealth self-management; however, we also identified 2 levels. On the basis of the explorative literature review, we discovered that depending on the user group (eg, older patients) or the mHealth app (eg, sensors with higher costs), price value played a significant role [18,61,63]. In contrast, the price was not relevant for less expensive mHealth apps such as smartphone apps [22,53,59]. We were also able to identify those aspects in the guided interviews, as shown in the extracts given in the following section.

Approximately all (7/8, 88%) of the mHealth users assigned a rather subordinate role to the price, especially for mHealth smartphone apps. Their focus was on the gain in convenience and quality of life. Health insurance companies usually cover the costs of the considerably more expensive CGM systems. However, the mHealth users agreed that even an appropriate copayment would not affect the use and acceptance of the system:

> *...so I do not have to pay anything for the smartphone app. I just need to pay for the sensors. So that is thirty euros a quarter, which is nothing. Even if the app had to be paid for, it depends on how much, of course...I would definitely pay...because it is a significant relief and would be worth it to me...* [User 2, female, type 1 diabetes]

If the price of an mHealth self-management app is within a reasonable range, most (9/11, 82%) of the mHealth or technology acceptance experts consider it to have no significant role. However, if the price is too high, it will affect acceptance, and people will not start or continue using the mHealth app:

> *...the price is often unimportant because the things are either free or paid for by the health insurance—so, in the very rarest cases, I have to spend a large amount of money for a specific application...* [Expert 3, male]

Habit

We were also able to confirm the importance of habit in diabetes mHealth self-management. On the basis of our explorative literature review, we identified several scientific articles that showed relevant aspects of habit in the context of mHealth self-management [53,59,62,64]. The authors emphasized the strong influence of habit on the expected outcome and the importance of continuous use because of regular patterns and routines. We were also able to identify those aspects in the guided interviews, as shown in the extracts given in the following section.

Approximately all (7/8, 88%) of the mHealth users stated that the mHealth app has taken an important place in their everyday life and has become a habit, improving their disease management:

> *...important place in my life. Compared to before, now with the app I test, I think, almost fifteen times more than before. It is already routine...I test much more than before with a standard test device...* [User 2, female, type 1 diabetes]

In addition, approximately half (5/11, 45%) of the mHealth or technology acceptance experts confirmed that integrating the mHealth app into daily routines is essential for sustained use:

> *...the habit is, of course, what drives you in the end, to do the same thing over and over again...* [Expert 4, male]

Additional Constructs

However, the analysis showed that additional constructs, as shown in Textbox 2, may also need to be considered to predict the user acceptance of mHealth self-management apps in diabetes.

Textbox 2. Newly proposed and confirmed constructs for the acceptance of mobile health self-management in diabetes.

Trust

- Degree of trust in the data collected by the mobile health app concerning data security, privacy, quality, and processing

Perceived disease threat

- Degree of patients' awareness of risks and limitations to health and well-being related to diabetes

Newly Proposed and Confirmed Construct: Trust

We identified the construct *trust* in several places in this study, showing its relevance to the field of mHealth self-management in diabetes. Trust can be defined as belief or confidence in other people or things [65]. We used the term *trust* to combine aspects such as data security, privacy, anonymity, and information quality. This approach is in line with a recent study on public trust in the health care system, in which the authors investigated various aspects that influence trust to understand the construct better [66]. On the basis of the explorative literature review, we identified several scientific articles that highlighted the importance and relevant aspects of trust in the context of mHealth self-management [18,21,22,40,59,65,67]. The authors emphasized the positive influence of trust as a crucial aspect in predicting acceptance and intention to use IT. We could also identify these aspects in the guided interviews, as shown in the extracts given in the following section.

Data protection and privacy were considered essential features of mHealth self-management apps by half (4/8, 50%) of the surveyed mHealth users. Problems with insufficient data protection and privacy can lead to low acceptance and termination of use:

> *...possibly lead to the fact that I stop...if I have the feeling my privacy is not maintained...* [User 1, female, type 2 diabetes]

In addition, from the perspective of most (10/11, 91%) of the mHealth or technology acceptance experts, data protection, data security, and privacy are central prerequisites for the acceptance and long-term use of mHealth apps. Specifically, the handling of data by third parties, such as service providers, has a significant influence on the use decision of mHealth apps:

> *...you also have to trust the app provider or manufacturer...if there is even the slightest risk that personal data is sold, and not anonymized at best...this does not increase trust, and the application probably will not be used...* [Expert 5, male]

Thus, according to more than half (7/11, 64%) of the surveyed mHealth or technology acceptance experts, it is not only about technical parameters of the mHealth app, such as data security. They increasingly see subjective factors such as trust in the service and service provider as relevant for mHealth acceptance:

> *...data protection is only one aspect...it is really about trust...[Expert 5, male]*

As the construct *trust* is not part of the UTAUT2 model but essential for accepting and using mHealth self-management apps, some (5/11, 45%) of the mHealth or technology acceptance experts recommended adding it:

> *...something like the trust that the data is not being misused...it is such a central aspect...because trust is, at least in Germany and I also think in Austria...a central component of consumer health IT applications.* [Expert 5, male]

Therefore, we were able to confirm the importance of trust for mHealth self-management in diabetes.

Newly Proposed and Confirmed Construct: Perceived Disease Threat

In this study, we identified the construct *perceived disease threat* in several places. The Health Belief Model first defined the construct of *perceived disease threat* [68,69]. The Health Belief Model refers to avoiding and preventing illness through specific health actions [70]. In this study, we used the construct of *perceived disease threat* to assess patients' awareness of risks and limitations to health and well-being associated with diabetes [70]. On the basis of our explorative literature review, we identified several scientific articles that highlight the importance and relevant aspects of perceived disease threat in the context of mHealth self-management [19,21,34,40,71]. In particular, the authors highlighted that patient awareness of the risks associated with chronic diseases could help to improve the acceptance of mHealth self-management apps. We were also able to identify those aspects in the guided interviews, as shown in the extracts given in the following section.

Many (5/8, 63%) of the mHealth users mentioned that they started to use a CGM system because of the negative impact on their blood glucose levels when they partially stopped using the conventional measurement because of the perceived

inconvenience of pricking their finger. In addition, the CGM system protects against dangerous situations such as nighttime hypoglycemia, which they highlighted to be essential for acceptance:

> *...the pricking was highly burdensome to me, so I partly stopped doing it, which was not really beneficial for developing blood glucose levels...* [User 3, male, type 2 diabetes]

> *...you never know what happens at night when you do not wake up when you have hypoglycemia, and if I didn't have the system, quite different things could happen...* [User 4, female, type 1 diabetes]

According to more than half (6/11, 55%) of the surveyed mHealth or technology acceptance experts, people who experience a disease and perceive it as a risk are more open to alternatives such as mHealth apps that promise positive benefits, which increases their acceptance:

> *...the patients' current state of health and suffering are essential...someone who has to ensure very extensive self-management is much more open-minded than someone who only has to collect or document data once a day or once a week...* [Expert 10, female]

Therefore, according to some (3/11, 27%) of the mHealth or technology acceptance experts, the UTAUT2 model should be extended with variables related to the disease state and the perceived disease threat:

> *...I would include disease-related variables...for example, chronic diseases...something like a perceived threat.* [Expert 9, female]

Therefore, we were able to confirm the importance of the perceived disease threat for mHealth self-management in diabetes.

Newly Proposed but Not Confirmed Construct: Personal Innovativeness

We were able to identify the construct of *personal innovativeness* only in the explorative literature review and the interviews with mHealth or technology acceptance experts but not with mHealth users. On the basis of the explorative literature review, we identified only 2 articles that highlighted the importance of personal innovativeness in the context of mHealth self-management [59,60]. The authors described personal innovativeness as the ability of a person to be open to new ideas and make innovative decisions [60]. As shown in the extract given in the following section, we were also able to identify the described aspect in the interviews with mHealth or technology acceptance experts.

Most (8/11, 73%) of the mHealth or technology acceptance experts see technology-savvy people and people who want to control their data as being particularly open to accepting and using mHealth apps:

> *...especially technically-savvy patients, as well as patients who do not want to travel to the hospital three times a week to record a certain value...* [Expert 6, male]

However, the construct is already part of the moderating effects of hedonic motivation on behavioral intention in the UTAUT2 model because of the associated differences in users' willingness to innovate [26]. Therefore, we were not able to confirm the importance of personal innovativeness as an additional construct for mHealth self-management in diabetes.

Discussion

Principal Findings

In this qualitative methods triangulation study, we used different perspectives to investigate whether the UTAUT2 model is suitable for predicting mHealth acceptance using diabetes as an example. Our results showed that we were able to confirm all exogenous UTAUT2 constructs. However, we verified that 2 essential constructs are missing in the UTAUT2 model to predict mHealth acceptance. We determined the constructs of *trust* and *perceived disease threat* to be relevant in this context. In contrast, the construct *personal innovativeness*, which we also identified, seemed less relevant for mHealth users, as we did not find indicators in the interviews. Furthermore, the construct *personal innovativeness* is already considered in the UTAUT2 model; therefore, it is unnecessary to add it as a separate construct.

Strengths and Limitations

We used a qualitative research method with its open approach to investigate the subject area.

The triangulation of explorative literature review (step 1) and guided interviews with mHealth or technology acceptance experts (step 2) and mHealth users (step 3) allowed us to identify relevant aspects influencing mHealth acceptance from different perspectives.

Using the method of structured content analysis combined with qualitative methods triangulation allowed us to confirm all relevant categories. In addition, we were able to identify less relevant categories; therefore, those categories are not required to be added to the UTAUT2 model. As expected, the systematic combination of the different methods proved successful, as we were able to confirm all exogenous UTAUT2 constructs and identify new categories quite clearly. The results have confirmed each other and can, therefore, be considered reliable.

We chose diabetes as an example as it is one of the most common chronic diseases for which mHealth apps are an essential element of self-management. Owing to the broad spectrum of patients with diabetes and available mHealth apps, the qualitative results also seem to be generalizable to mHealth apps for other chronic diseases.

We systematically selected different interview participants and triangulated different sources of information. We also followed the principle of theoretical saturation and are, therefore, confident that we have captured all relevant aspects. However, as the selection of mHealth users focused on active users, there might have been some selection bias.

Although the interviews were only conducted with people from Austria and Germany using diabetes as an example, we consider that the qualitative methods triangulation study results also

apply to countries with comparable health care systems, technical infrastructure, socioeconomic and cultural backgrounds, and other chronic diseases where mHealth self-management apps are used because of the multicenter study design.

A risk when conducting interviews is that people's responses may be influenced by social desirability. We tried to reduce this risk by creating a trustful and open interview atmosphere in which only the interviewer and the interview candidate were present.

Although we have adhered to the quality criteria of qualitative research concerning objectivity, reliability, and validity by applying neutrality in data analysis, rule guidance in the research process, peer debriefing, and method triangulation, explorative studies are associated with certain limitations such as generalizability. Therefore, we plan to verify the results within the framework of a quantitative follow-up study.

Comparison With Prior Work

To date, there have not been many studies that have used the UTAUT2 model to predict mHealth acceptance [41,43,44,60,62,64]. In addition, only a few of these studies have explicitly highlighted the suitability of the UTAUT2 model in this context [43,60,64]. With our qualitative methods triangulation study, we were also able to confirm the suitability of the UTAUT2 model for predicting mHealth acceptance.

In our results, we showed that the four exogenous UTAUT constructs of *facilitating conditions*, *performance expectancy*, *effort expectancy*, and *social influence* are relevant to the acceptance of mHealth in diabetes, which is consistent with previous mHealth studies [19,36,37,41].

We were also able to verify the relevance of the three additional exogenous UTAUT2 constructs: *hedonic motivation*, *price value*, and *habit*. In particular, *hedonic motivation* and *habit* were highlighted to be essential for the acceptance and long-term use of mHealth self-management apps in diabetes. In their study, the authors pointed out the importance of both constructs for mHealth acceptance [64].

Our results showed that the price of an mHealth self-management app is considered less relevant by mHealth or technology acceptance experts and mHealth users, who focus more on the benefits of the app. This observation is consistent with the findings from previous studies, where the authors showed that price value does not influence mHealth acceptance [43,59].

In addition to the exogenous UTAUT2 constructs, we identified three relevant constructs: *trust*, *perceived disease threat*, and *personal innovativeness*. The relevance of *trust* and *perceived disease threat* were highlighted in our results as essential aspects for mHealth acceptance in diabetes. This observation aligns with previous studies where the authors described the relevance of *trust* in adopting different eHealth services by extending the TAM and UTAUT model [65,67].

The relevance of the construct *perceived disease threat* was also confirmed by several studies where the authors used the TAM and UTAUT model to investigate the acceptance and adoption of mHealth apps in patients with chronic diseases such as hypertension and diabetes [19,21,34,71].

However, our results showed that the construct of *personal innovativeness* turned out to be less relevant. This observation is consistent with the original UTAUT2 study in which the authors described the construct *personal innovativeness* as an implicit moderating effect of the construct *hedonic motivation* on *behavioral intention* [26].

Conclusions

In summary, our study showed that the UTAUT2 model is suitable for predicting mHealth acceptance, as shown in the field of mHealth for diabetes. However, we also showed that the additional constructs of *trust* and *perceived disease threat* are required to comprehensively examine mHealth acceptance in this context.

We see great potential for an extended UTAUT2 model that focuses on additional mHealth predictors. Further research is needed to determine whether the newly identified constructs also apply to other mHealth apps and clinical settings.

Acknowledgments

The authors would like to gratefully acknowledge the Austrian Diabetes Association for their help with patient recruitment. In addition, they would like to thank all the participants in the interviews for sharing their experiences and knowledge with them.

Conflicts of Interest

None declared.

References

1. World Health Organization. Global Report on Diabetes. Geneva: World Health Organization; 2016.
2. Hood M, Wilson R, Corsica J, Bradley L, Chirinos D, Vivo A. What do we know about mobile applications for diabetes self-management? A review of reviews. J Behav Med 2016 Dec;39(6):981-994. [doi: 10.1007/s10865-016-9765-3] [Medline: 27412774]
3. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]

4. Cappon G, Vettoretti M, Sparacino G, Facchinetti A. Continuous glucose monitoring sensors for diabetes management: a review of technologies and applications. Diabetes Metab J 2019 Aug;43(4):383-397 [FREE Full text] [doi: 10.4093/dmj.2019.0121] [Medline: 31441246]
5. Jeffrey B, Bagala M, Creighton A, Leavey T, Nicholls S, Wood C, et al. Mobile phone applications and their use in the self-management of Type 2 Diabetes Mellitus: a qualitative study among app users and non-app users. Diabetol Metab Syndr 2019;11:84 [FREE Full text] [doi: 10.1186/s13098-019-0480-4] [Medline: 31636719]
6. Fu H, McMahon SK, Gross CR, Adam TJ, Wyman JF. Usability and clinical efficacy of diabetes mobile applications for adults with type 2 diabetes: a systematic review. Diabetes Res Clin Pract 2017 Sep;131:70-81. [doi: 10.1016/j.diabres.2017.06.016] [Medline: 28692830]
7. Veazie S, Winchell K, Gilbert J, Paynter R, Ivlev I, Eden KB, et al. Rapid evidence review of mobile applications for self-management of diabetes. J Gen Intern Med 2018 Jul;33(7):1167-1176. [doi: 10.1007/s11606-018-4410-1] [Medline: 29740786]
8. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: a systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018 Dec;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]
9. Slater H, Campbell JM, Stinson JN, Burley MM, Briggs AM. End user and implementer experiences of mHealth technologies for noncommunicable chronic disease management in young adults: systematic review. J Med Internet Res 2017 Dec 12;19(12):e406 [FREE Full text] [doi: 10.2196/jmir.8888] [Medline: 29233804]
10. Delva S, Mendez KJ, Cajita M, Koirala B, Shan R, Wongvibulsin S, et al. Efficacy of mobile health for self-management of cardiometabolic risk factors: a theory-guided systematic review. J Cardiovasc Nurs 2021;36(1):34-55 [FREE Full text] [doi: 10.1097/JCN.0000000000000659] [Medline: 32040072]
11. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 Apr;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
12. Trawley S, Baptista S, Browne JL, Pouwer F, Speight J. The use of mobile applications among adults with type 1 and type 2 diabetes: results from the Second MILES-Australia (MILES-2) study. Diabetes Technol Ther 2017 Dec;19(12):730-738. [doi: 10.1089/dia.2017.0235] [Medline: 29028442]
13. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
14. Liang Z, Nagata Y, Martell M, Nishimura T. Nurturing wearable and mHealth technologies for self-care: mindset, tool set and skill set. In: Proceedings of the IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom). 2016 Presented at: IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom); Sept. 14-16, 2016; Munich, Germany. [doi: 10.1109/healthcom.2016.7749432]
15. Canhoto AI, Arp S. Exploring the factors that support adoption and sustained use of health and fitness wearables. J Mark Manag 2016 Oct 26;33(1-2):32-60. [doi: 10.1080/0267257X.2016.1234505]
16. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
17. Dillon A. User Acceptance of Information Technology. London: Taylor and Francis; 2001.
18. Scheibe M, Reichelt J, Bellmann M, Kirch W. Acceptance factors of mobile apps for diabetes by patients aged 50 or older: a qualitative study. Med 2.0 2015 Mar 02;4(1):e1 [FREE Full text] [doi: 10.2196/med20.3912] [Medline: 25733033]
19. Breil B, Kremer L, Hennemann S, Apolinário-Hagen J. Acceptance of mHealth apps for self-management among people with hypertension. Stud Health Technol Inform 2019 Sep 03;267:282-288. [doi: 10.3233/SHTI190839] [Medline: 31483283]
20. Pais S, Parry D, Petrova K, Rowan J. Acceptance of using an ecosystem of mobile apps for use in diabetes clinic for self-management of gestational diabetes mellitus. Stud Health Technol Inform 2017;245:188-192. [Medline: 29295079]
21. Zhang Y, Liu C, Luo S, Xie Y, Liu F, Li X, et al. Factors influencing patients' intentions to use diabetes management apps based on an extended unified theory of acceptance and use of technology model: web-based survey. J Med Internet Res 2019 Aug 13;21(8):e15023 [FREE Full text] [doi: 10.2196/15023] [Medline: 31411146]
22. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
23. Apolinário-Hagen J, Menzel M, Hennemann S, Salewski C. Acceptance of mobile health apps for disease management among people with multiple sclerosis: web-based survey study. JMIR Form Res 2018 Dec 12;2(2):e11977 [FREE Full text] [doi: 10.2196/11977] [Medline: 30684408]
24. Zhang Y, Li X, Luo S, Liu C, Xie Y, Guo J, et al. Use, perspectives, and attitudes regarding diabetes management mobile apps among diabetes patients and diabetologists in china: national web-based survey. JMIR Mhealth Uhealth 2019 Feb 08;7(2):e12658 [FREE Full text] [doi: 10.2196/12658] [Medline: 30735147]
25. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man-Machine Stud 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
26. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]

27. International Diabetes Federation. IDF Diabetes Atlas: Ninth Edition. Brussels, Belgium: International Diabetes Federation; 2019.
28. Ammenwerth E. Technology acceptance models in health informatics: TAM and UTAUT. Stud Health Technol Inform 2019 Jul 30;263:64-71. [doi: 10.3233/SHTI190111] [Medline: 31411153]
29. Rahimi B, Nadri H, Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Dec;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
30. 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]
31. 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]
32. Cho J. The impact of post-adoption beliefs on the continued use of health apps. Int J Med Inform 2016 Mar;87:75-83. [doi: 10.1016/j.ijmedinf.2015.12.016] [Medline: 26806714]
33. Jeon E, Park H. Factors affecting acceptance of smartphone application for management of obesity. Healthc Inform Res 2015 Apr;21(2):74-82 [FREE Full text] [doi: 10.4258/hir.2015.21.2.74] [Medline: 25995959]
34. Zhu Z, Liu Y, Che X, Chen X. Moderating factors influencing adoption of a mobile chronic disease management system in China. Inform Health Soc Care 2018 Jan;43(1):22-41. [doi: 10.1080/17538157.2016.1255631] [Medline: 28068149]
35. Rönnby S, Lundberg O, Fagher K, Jacobsson J, Tillander B, Gauffin H, Dahlström, et al. mHealth self-report monitoring in competitive middle- and long-distance runners: qualitative study of long-term use intentions using the technology acceptance model. JMIR Mhealth Uhealth 2018 Aug 13;6(8):e10270 [FREE Full text] [doi: 10.2196/10270] [Medline: 30104183]
36. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
37. Quaosar GM, Hoque MR, Bao Y. Investigating factors affecting elderly's intention to use m-health services: an empirical study. Telemed J E Health 2018 Apr;24(4):309-314. [doi: 10.1089/tmj.2017.0111] [Medline: 28976824]
38. Lee S, Choi MJ, Rho MJ, Kim D, Choi IY. Factors affecting user acceptance in overuse of smartphones in mobile health services: an empirical study testing a modified integrated model in South Korea. Front Psychiatry 2018;9:658 [FREE Full text] [doi: 10.3389/fpsyt.2018.00658] [Medline: 30631283]
39. Ma Q, Chan AH, Chen K. Personal and other factors affecting acceptance of smartphone technology by older Chinese adults. Appl Ergon 2016 May;54:62-71. [doi: 10.1016/j.apergo.2015.11.015] [Medline: 26851465]
40. Peng W, Yuan S, Holtz BE. Exploring the challenges and opportunities of health mobile apps for individuals with type 2 diabetes living in rural communities. Telemed J E Health 2016 Sep;22(9):733-738. [doi: 10.1089/tmj.2015.0180] [Medline: 26982017]
41. Alam MM, Alam MZ, Rahman SA, Taghizadeh SK. Factors influencing mHealth adoption and its impact on mental well-being during COVID-19 pandemic: a SEM-ANN approach. J Biomed Inform 2021 Apr;116:103722. [doi: 10.1016/j.jbi.2021.103722] [Medline: 33705856]
42. Sitar-Tăut D. Mobile learning acceptance in social distancing during the COVID-19 outbreak: the mediation effect of hedonic motivation. Hum Behav Emerg Technol 2021 May 24 [FREE Full text] [doi: 10.1002/hbe2.261] [Medline: 34222833]
43. Salgado T, Tavares J, Oliveira T. Drivers of mobile health acceptance and use from the patient perspective: survey study and quantitative model development. JMIR Mhealth Uhealth 2020 Jul 09;8(7):e17588 [FREE Full text] [doi: 10.2196/17588] [Medline: 32673249]
44. Tomczyk S, Barth S, Schmidt S, Muehlan H. Utilizing health behavior change and technology acceptance models to predict the adoption of COVID-19 contact tracing apps: cross-sectional survey study. J Med Internet Res 2021 May 19;23(5):e25447 [FREE Full text] [doi: 10.2196/25447] [Medline: 33882016]
45. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
46. Kuckartz U. Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. 4. Auflage. Weinheim, Basel: Beltz Juventa; 2018.
47. Saldaña J. The Coding Manual for Qualitative Researchers. Thousand Oaks, CA: SAGE Publications; 2016.
48. Misoch S. Qualitative Interviews. 2., Erweiterte Und Aktualisierte Auflage. Berlin: De Gruyter Oldenbourg; 2019.
49. Kujala S, Roto V, Väänänen-Vainio-Mattila K, Karapanos E, Sinnelä A. UX Curve: a method for evaluating long-term user experience. Interact Comput 2011 Sep;23(5):473-483. [doi: 10.1016/j.intcom.2011.06.005]
50. Rädiker S, Kuckartz U. Analyse Aualitativer Daten Mit MAXQDA: Text, Audio und Video. Wiesbaden: Springer; 2019.
51. Flick U. Triangulationine Einführung. 2nd ed. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden; 2008.
52. Kuckartz U. Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren. Wiesbaden: Springer; 2014.
53. Chang C, Lee T, Mills ME. Experience of home telehealth technology in older patients with diabetes. Comput Inform Nurs 2017 Oct;35(10):530-537. [doi: 10.1097/CIN.0000000000000341] [Medline: 28291156]

54. Humble JR, Tolley EA, Krukowski RA, Womack CR, Motley TS, Bailey JE. Use of and interest in mobile health for diabetes self-care in vulnerable populations. J Telemed Telecare 2016 Jan;22(1):32-38. [doi: 10.1177/1357633X15586641] [Medline: 26026179]
55. Lee J, Rho MJ. Perception of influencing factors on acceptance of mobile health monitoring service: a comparison between users and non-users. Healthc Inform Res 2013 Sep;19(3):167-176 [FREE Full text] [doi: 10.4258/hir.2013.19.3.167] [Medline: 24175115]
56. Nelson LA, Mulvaney SA, Johnson KB, Osborn CY. mHealth intervention elements and user characteristics determine utility: a mixed-methods analysis. Diabetes Technol Ther 2017 Jan;19(1):9-17. [doi: 10.1089/dia.2016.0294] [Medline: 28099052]
57. Torbjørnsen A, Ribu L, Rønnevig M, Grøttland A, Helseth S. Users' acceptability of a mobile application for persons with type 2 diabetes: a qualitative study. BMC Health Serv Res 2019 Sep 06;19(1):641 [FREE Full text] [doi: 10.1186/s12913-019-4486-2] [Medline: 31492176]
58. Torbjørnsen A, Småstuen MC, Jenum AK, Årsand E, Ribu L. Acceptability of an mHealth app intervention for persons with type 2 diabetes and its associations with initial self-management: randomized controlled trial. JMIR Mhealth Uhealth 2018 May 21;6(5):e125 [FREE Full text] [doi: 10.2196/mhealth.8824] [Medline: 29784635]
59. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
60. Huang C, Yang M. Empirical investigation of factors influencing consumer intention to use an artificial intelligence-powered mobile application for weight loss and health management. Telemed J E Health 2020 Oct;26(10):1240-1251. [doi: 10.1089/tmj.2019.0182] [Medline: 31971883]
61. Macdonald EM, Perrin BM, Hyett N, Kingsley MI. Factors influencing behavioural intention to use a smart shoe insole in regionally based adults with diabetes: a mixed methods study. J Foot Ankle Res 2019;12:29 [FREE Full text] [doi: 10.1186/s13047-019-0340-3] [Medline: 31139261]
62. Yuan S, Ma W, Kanthawala S, Peng W. Keep using my health apps: discover users' perception of health and fitness apps with the UTAUT2 model. Telemed J E Health 2015 Apr 28:735-741. [doi: 10.1089/tmj.2014.0148] [Medline: 25919238]
63. de Veer AJ, Peeters JM, Brabers AE, Schellevis FG, Rademakers JJ, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15:103 [FREE Full text] [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
64. Duarte P, Pinho JC. A mixed methods UTAUT2-based approach to assess mobile health adoption. J Busin Res 2019 Sep;102:140-150. [doi: 10.1016/j.jbusres.2019.05.022]
65. Alaiad A, Zhou L. The determinants of home healthcare robots adoption: an empirical investigation. Int J Med Inform 2014 Nov;83(11):825-840. [doi: 10.1016/j.ijmedinf.2014.07.003] [Medline: 25132284]
66. Gille F, Smith S, Mays N. What is public trust in the healthcare system? A new conceptual framework developed from qualitative data in England. Soc Theory Health 2020 Jan 29;19(1):1-20. [doi: 10.1057/s41285-020-00129-x]
67. Hoque MR, Bao Y, Sorwar G. Investigating factors influencing the adoption of e-Health in developing countries: a patient's perspective. Inform Health Soc Care 2017 Jan;42(1):1-17. [doi: 10.3109/17538157.2015.1075541] [Medline: 26865037]
68. Becker MH, Drachman RH, Kirscht JP. A new approach to explaining sick-role behavior in low-income populations. Am J Public Health 1974 Mar;64(3):205-216. [doi: 10.2105/ajph.64.3.205] [Medline: 4811762]
69. Rosenstock IM. Historical origins of the health belief model. Health Edu Monogr 1974 Dec 01;2(4):328-335. [doi: 10.1177/109019817400200403]
70. Janz NK, Becker MH. The health belief model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
71. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' acceptance of smartphone health technology for chronic disease management: a theoretical model and empirical test. JMIR Mhealth Uhealth 2017 Dec 06;5(12):e177 [FREE Full text] [doi: 10.2196/mhealth.7886] [Medline: 29212629]

Abbreviations

CGM: continuous glucose monitoring
COREQ: Consolidated Criteria for Reporting Qualitative Research
IT: information technology
mHealth: mobile health
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology
UTAUT2: Unified Theory of Acceptance and Use of Technology 2
UX: user experience

Edited by A Kushniruk; submitted 12.11.21; peer-reviewed by Z Huang; comments to author 06.01.22; revised version received 27.01.22; accepted 03.02.22; published 09.03.22.

Please cite as:
Schretzlmaier P, Hecker A, Ammenwerth E
Suitability of the Unified Theory of Acceptance and Use of Technology 2 Model for Predicting mHealth Acceptance Using Diabetes as an Example: Qualitative Methods Triangulation Study
JMIR Hum Factors 2022;9(1):e34918
URL: https://humanfactors.jmir.org/2022/1/e34918
doi:10.2196/34918
PMID:35262493

Original Paper

Patients' Experiences of Using a Smartphone App After Cardiac Rehabilitation: Qualitative Study

Pernille Lunde[1], PhD; Asta Bye[2,3], PhD; Kari Anette Bruusgaard[1], MPH; Elisabet Hellem[1], MSc; Birgitta Blakstad Nilsson[1,4], PhD

[1]Department of Physiotherapy, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway

[2]Department of Nursing and Health Promotion, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway

[3]European Palliative Care Research Centre, Department of Oncology, Oslo University Hospital, Oslo, Norway

[4]Section for Physiotherapy, Division of Medicine, Oslo University Hospital, Oslo, Norway

Corresponding Author:
Pernille Lunde, PhD
Department of Physiotherapy
Faculty of Health Sciences
Oslo Metropolitan University
PB 4
St. Olavs Plass
Oslo, 0130
Norway
Phone: 47 48063537
Email: plunde@oslomet.no

Abstract

Background: Exercise-based cardiac rehabilitation (CR) is a crucial part of the treatment of patients with cardiac diseases, and adherence to healthy behavior is a prerequisite to improve long-term prognosis. Unfortunately, adherence to healthy behavior adapted in CR is challenging for many cardiac patients in the long term. Recently, we demonstrated that follow-up conducted via an app for 1 year significantly improved adherence to healthy behavior after CR. To increase the knowledge and understanding of mobile Health (mHealth) interventions that can promote acceptance and adherence, qualitative research investigating patients' experiences with these interventions is warranted.

Objective: The aim was to investigate patient experiences with individualized long-term follow-up conducted via an app for 1 year and their thoughts about what features promoted adherence to healthy behavior after CR. The purpose was to increase the understanding of significant findings previously reported and to guide future development of similar interventions in the field of adherence.

Methods: A qualitative study with individual interviews was conducted from November 2018 to May 2019. A thematic interview guide was used when conducting the semistructured in-depth interviews. The interviews were audio recorded and transcribed successively during the period in which the interviews were conducted. Texts were managed and systematized by NVivo. Interviews were analyzed by qualitative content analysis. Codes and themes were inductively developed.

Results: Ten patients who had participated in a randomized controlled trial evaluating the effect of follow-up conducted via an app on adherence to healthy behavior after CR were included. The median patient age was 65 years (range 46-72 years), and both genders were represented. The analysis resulted in the following 4 themes describing the patients' experiences: (1) The person behind the app is crucial for motivation and adherence; (2) The app as a commitment; (3) The app as a path to independence; and (4) Suggestions for improvements. Features experienced as beneficial to promote adherence were individualized feedback and the use of goal setting. The significance of the person behind the app (the supervisor) who provided individualized feedback was a consistent finding. This person seemed to promote motivation in general and to enable other known behavioral change techniques.

Conclusions: The person behind the app (the supervisor) seems to be one of the most significant success factors in promoting adherence to healthy behavior after CR. This indicates that a health care provider must actively participate in a patient's process of adherence to healthy behavior, even when using interventions, including an app. Future development of interventions in the field of adherence should strive to create tools that enable an ongoing collaborative relationship between the patient and the health care provider. The follow-up should be based on the patient's own goals, and individualized feedback should be provided.

(JMIR Hum Factors 2022;9(1):e34294) doi:10.2196/34294

KEYWORDS

mHealth; mobile health; cardiac rehabilitation; mobile phone app; smartphone; lifestyle

Introduction

Exercise-based cardiac rehabilitation (CR) is a crucial part of the treatment of patients with cardiac diseases and is a Class IA recommendation in European guidelines [1,2]. The overall goal of secondary prevention, including CR, is to prevent subsequent cardiac events [2,3]. In this context, adherence to healthy behavior, including physical activity, regular exercise, healthy nutrition with bodyweight control, compliance with taking medication, and smoking cessation [2], is crucial. Although adherence to healthy behavior is a prerequisite to improve the long-term prognosis, the majority of cardiac patients do not achieve the guideline standard for secondary prevention in the long term [2,4]. Research evaluating interventions aiming to improve adherence to healthy behavior after CR is therefore warranted [2].

Mobile health (mHealth) interventions have been proposed to meet the challenges related to adherence to healthy behavior and have thus been suggested as potential interventions after CR [2,5-7]. In particular, smartphone apps have been considered promising owing to their ability to monitor patients' health from anywhere at any time [5,8]. Previous research has highlighted the need for individualization of such interventions [9,10]. Recently, we demonstrated the feasibility of using an app to provide individualized follow-up in patients after CR [11]. Based on the results from this study, we developed and conducted a randomized controlled trial (RCT) aiming to evaluate the effect of individualized follow-up with an app for 1 year on health outcomes relevant for adherence to healthy behavior in patients after CR [12]. Patients in the intervention group received access to an app where they added individual goals and accompanying tasks [12,13]. They were monitored and followed by a supervisor (specialized physiotherapist) for a year. The app itself provided reminders and evaluations of tasks and weekly goal achievement, and the patients could write notes related to each goal. The intervention included comprehensive individualized feedback, based on the patients' goals and what they had done, through an email every week for the first 12 weeks and every fourth week for the rest of the year. Throughout the year, they also received between 1 and 3 short motivational messages every week. These messages were written individually for each patient. However, sometimes the content was of a more general nature. Additionally, patients could submit questions to the supervisor and receive answers within 2 working days throughout the year [12,13]. The results demonstrated that using the app significantly improved peak oxygen uptake, exercise performance, exercise habits, and self-perceived goal achievement, compared with a control group that received usual care after CR [12]. All patients allocated to the intervention group used the app, and as much as 71% of the patients used the app on a daily or weekly basis throughout the year [12].

The high acceptance and use of the app in our study was unique as difficulty or low acceptance in using the technology is a frequent obstacle in similar interventions [14]. In order to increase the knowledge and understanding of mHealth functions and components that can promote acceptance and adherence, qualitative research investigating patients' experiences with these interventions is urged. To our knowledge, no previous studies have explored patients' experiences with individualized mHealth interventions lasting for a whole year. The purpose of this study was to increase the understanding of the significant findings previously reported [12] and to guide future development of similar interventions in the field of adherence. Our aim was to investigate patient experiences with individualized long-term follow-up conducted via an app for 1 year, in order to gain more knowledge about features promoting adherence to healthy behavior after CR.

Methods

Design

A qualitative study with individual interviews was conducted to describe patients' experiences with a long-term follow-up intervention conducted via an app. The interviews were planned to be completed within 2 weeks after the patients had ended their follow-up period of 1 year in the previously mentioned RCT [12].

Recruitment

Participants were recruited from the RCT [12] (n=113). Enrollment in the RCT was carried out at 2 CR centers in the eastern part of Norway from October 2017 to June 2018. These CR centers offered, in total, 3 different CR programs: 12-week outpatient CR, 4-week inpatient CR, and 1-week inpatient CR. The randomization was stratified by the CR program to ensure equal participation and thereby representativeness.

At the 1-year follow-up assessment, participants in the intervention group were recruited in this study. Living nearby Oslo (maximum 1-hour commute) was set as an inclusion criterion as the interviews were planned to be conducted at Oslo Metropolitan University (OsloMet) in Oslo, Norway. Efforts were made to ensure that participants were representative of the CR population in the eastern part of Norway (both genders, participation in different CR programs, and different ages). Recruitment and inclusion in this study continued until data saturation was achieved.

Interviews and Interview Guide

Individual interviews were completed from November 2018 to May 2019. The interviews were carried out at OsloMet. One participant chose to complete the interview digitally owing to several unforeseen appointments that made it difficult to attend physically. To ensure sufficient quality on the audio recording, Skype for Business was used. The interviews lasted from 35 to 62 minutes (44 minutes on average) and were carried out by 2

researchers (KAB and EH) who did not take part in the RCT from which the participants were recruited. Both interviewers had extensive experience from CR and qualitative research. To ensure the material was as comprehensive as possible, all interviews were carried out with both researchers present.

A thematic interview guide (Multimedia Appendix 1) was used when conducting the semistructured in-depth interviews. The interview guide was developed by 3 of the authors (PL, KAB, and EH) and validated by all authors. The aim was to maintain an open nonjudgmental attitude. Emphasis was placed on listening to the responses to open-ended questions and allowing the participants to fully explain a phenomenon, together with an invitation to reflect upon their experiences [15]. The interviews were audio recorded and transcribed successively during the period in which the interviews were conducted. Initially, the texts were managed and systematized in Microsoft Word, and by working manually with printouts and pen and paper. Thereafter, the texts were imported to, and managed and systematized by NVivo (released in March 2020) [16]. Quotations from the texts were translated from Norwegian to English by the first author (PL) and then validated by all co-authors.

Data Analysis

Transcribed interviews were analyzed by a thematic coding technique based on the framework by Braun and Clarke [17], which is a method for identifying, analyzing, and reporting patterns within qualitative data. The method includes the following 6 phases: (1) familiarization with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) writing the report [17]. The codes and themes were inductively developed.

Initially, the analysis involved repeated readings of each transcript by all the authors to obtain an overall impression of the material. The next phase involved coding the entire data set on a semantic level. Specifically, we focused on the parts of the data that revealed relevant information and descriptions regarding the current overall research question. Further, codes that revealed similar aspects of the data were grouped into preliminary themes, which were checked for consistency and variability within and across interviews. Subsequently, we identified and interpreted 4 overarching themes in a constant process of moving between the data, potential themes, and maps made for visualization, as well as in reference to relevant literature, and discussions and mutual understanding among the authors. Finally, themes were established if they were coherent and represented the meanings found in the interviews [17]. Throughout the analytic process, all findings were discussed and validated within the research group. In case of inconsistencies, further discussions and reflections were used for resolution.

Ethical Considerations

This study was approved by the Regional Committee for Medical and Health Research Ethics (South-East ID: 2016-1476) as a substudy of the previously described RCT. All included patients provided written informed consent.

Results

General Findings

Ten patients with a median age of 65 years (range 46-72 years) participated in the study (Table 1). More than half of the patients were retired. The majority had participated in a 4-week inpatient CR program or a 12-week outpatient CR program before inclusion in the RCT. All patients had their own goals or tasks related to exercise and physical activity. Additionally, 7 of the patients had goals related to weight loss or maintenance of bodyweight, and accompanying tasks were specific nutritional advice learned or implemented in primary CR. The numbers and types of goals are presented in Table 1. Nine of the patients attended the interview as scheduled, while 1 was unable to attend until 4 weeks after the follow-up assessment, due to other medical and social appointments.

All patients in the study mentioned that they used the app for preventive activities, such as exercise, physical activity, and healthy nutrition, and, without exception, they found the app easy to use. The patients' experiences evolved within the following 4 themes: (1) The person behind the app is crucial for motivation and adherence; (2) The app as a commitment; (3) The app as a path to independence; and (4) Suggestions for improvements. The first overarching theme was abstracted to subthemes.

Table 1. Characteristics of the patients (N=10).

Characteristic	Value, n (%)
Gender	
Male	9 (90)
Female	1 (10)
Age distribution (years)	
40-49	1 (10)
50-59	2 (20)
60-69	5 (50)
70-79	2 (20)
Civil status	
Married/cohabiting	8 (80)
Single	2 (20)
Employment status	
Employed	3 (30)
Retired	6 (60)
Disability benefits	1 (10)
Disease	
Coronary artery disease	7 (70)
Valve surgery	3 (30)
Type of cardiac rehabilitation	
One week	1 (10)
Four weeks	4 (40)
Twelve weeks	5 (50)
Smartphone	
iPhone	7 (70)
Android	3 (30)
Number of goals	
One	4 (40)
Two	6 (60)
Type of goal	
Exercise-related goal	9 (90)
Weight loss/maintenance goal	7 (70)

The Person Behind the App is Crucial for Motivation and Adherence

All patients in the study highlighted that the person behind the app (the supervisor) was considered a prerequisite to succeed with the intervention. However, this person cannot be just anyone. The patients highlighted that the person must possess a set of characteristics that primarily helps create a relationship of trust between the supervisor and the patient, which helps to make the app motivating and thereby helps the patient adhere to healthy behavior. Personal characteristics of special importance included engagement, professional competence, care, and support.

> *You know, she is not just anyone, the fact is that she gets involved and shows care and engagement in me as a person. At least I perceive it as if she wants my best, and she gives the advice that is for my best.* [Participant #1]

> *It is about the individual behind it, from whom you can almost experience a kind of love, and a person who is engaged in you.* [Participant #10]

The following 3 subthemes evolved from this theme: (1) individualized feedback, (2) follow-up based on own goals, and (3) a lifebuoy in the event of unforeseen events. The person behind the app was the common denominator for all 3 subthemes.

Individualized Feedback

An important motivating factor, which was highlighted by all patients in this study, was the individual feedback that each one received throughout the study period. The person behind the app made it possible to provide tailored feedback, advice, and guidance, which seems to have been a success criterion. The tailoring should be based on the patients' individual condition, the recent development, and the patients' likes and dislikes. This reinforces the feeling that the feedback is directed at the individual and not in general. Several of the patients used other general health apps during the intervention period. They pointed out the difference between individual feedback and automated feedback.

> *I do not really believe in apps providing feedback automatically. So, this app is great because there is a person providing the feedback, which means that the feedback is directed solely to you. That is, I think that is crucial, because this is what's motivates me.* [Participant #8]

> *The fact that there was a physical person, that you actually knew at the other end, who provided individualized feedback and you had the opportunity to communicate with, was extra motivating. This made it easier to keep up the good work.* [Participant #7]

Follow-up Based on Own Goals

The person behind the app enabled the follow-up to be based on individualized goals, which most patients highlighted as important to increase motivation.

> *I think that's pretty essential (setting your own goals). Of course, the more you personify this, the better it is. And of course, following them then. However, those goals could have been more nuanced. Maybe there could have been a few more.* [Participant #3]

A Lifebuoy in the Event of Unforeseen Events

Some of the patients experienced dramatic events that caused a significant setback during the year of follow-up. They expressed that for them, most likely, the app and follow-up had been extra important for long-term adherence to healthy behavior. In these cases, the patients mentioned that it was absolutely crucial that there was a person behind the app with whom they had an established and trustful relationship.

> *If I didn't have the app, or should I say "her"….. If I didn't have her at that time, I think I would have had extensive challenges getting to where I am today…, so fast... I would probably have walked and strolled a bit, but I would not have been able to physically be where I am today. Because of that (setback), I needed help in a proper way… not like "you have to do this, and you have to do that." …But something motivating and encouraging, and that is exactly what I got from her.* [Participant #4]

Additionally, adjustments and flexibility in goal-setting processes and the accompanying tasks were highlighted as central. This seems to be particularly important following dramatic events, when patients often must take one day at a time.

> *For me it has certainly had an extra great significance, because it was a bit like a crisis, and she came out with suggestions for alternatives to the goals I had set myself.* [Participant #1]

The App as a Commitment

Several of the patients described that the app, and the follow-up, provided a form of commitment. The commitment to the person behind the app turned out to be the most evident.

> *To be honest, I did not want to disappoint the supervisor, because she had been so motivating. So, my wife said, "It doesn't matter what I say, but when she says it, then it is important." So, maybe there`s something in it.* [Participant #4]

Several of the patients also expressed a commitment to the research project as a motivating factor. However, the distinction between the research project and the person behind the app was not clear.

> *We also knew that we were part of a research project, so you kind of felt it was a bit important what you were doing. Or at least, it could make a difference to her work. That she was involved, and that it was fun to try to take it seriously…. and then, the idea with that app and the follow-up was that you should perform at your best level …. that was a good motivation.* [Participant #10]

Finally, the patients expressed that the app also gave a commitment to themselves. To be challenged at their individual level was highlighted as motivating. Some patients described that they used the note function in the app and wrote a diary to give themselves an extra challenge, beyond the one they received from the supervisor.

> *I posted such a summary, that this week I have completed 4×4 intervals, while this week I have had pyramid intervals […]. So, it was a small summary for each week, and I really appreciated it because it was very nice to be able to scroll through, and it gave a motivation to keep up the good work and to challenge myself. It also gave me bad conscience if I did not exercise enough.* [Participant #7]

The App as a Path to Independence

The patients expressed that they experienced the downward adjustment regarding frequency of comprehensive feedback, also known as individualized feedback, as overwhelming and a bit scary. Despite this, the downgrading was perceived as important to increase independence while they at the same time felt safe and supported on a regular basis. Additionally, they knew they could easily get in touch with the supervisor if needed.

> *Right away it was a little shocking, like "Oh? Is it only once a month, now?" It was so nice with that attention…. But then, sort of, yes, that was the deal. I have reached a higher level…. Now, I must be more*

> *independent. […] I must keep it going by myself, so in that sense it gives a natural transition. But I felt like I had been living in a suite, a first-class suite, and then suddenly, I was down to third class, sort of.* [Participant #10]

Most of the patients expressed that the feeling of safety that the app gave them was important to promote and push themselves to the activities that they needed to reach their goals. In particular, when the frequency of the comprehensive feedback was downgraded, this safety was extra important. Through the 9 months with less frequent follow-up, they got the chance to experience that they were able to adhere to their program almost by themselves.

> *So, at that time when the frequency was downgraded, I was a bit alone. However, with that app and follow-up, you have a direct link to the expertise in a way, which is both reassuring and motivating. […] It was a very good safety net, it`s like wearing a parachute. You don't have to use it, but you know it`s there.* [Participant #7]

Suggestions for Improvements

Despite the promising result in the RCT regarding the effect of follow-up with the app, we also analyzed the qualitative data to illuminate the potential for improvements to optimize future interventions in the field of adherence. Overall, patients expressed high satisfaction with the app and justified this with the fact that it was easy to use. Most of the patients considered the app to be a tool, enabling human interaction.

> *So, technology can't replace people, but it is a helpful tool* [Participant #1]

Nevertheless, 2 suggestions for improvements clearly evolved. This was related to ownership of own goals and self-perceived goal achievement. Although most of the patients found it both meaningful and motivating that the use of the app and the follow-up were based on their own goals, ownership to some of the patients' goals could be questioned. Some of the patients expressed that their goals were made by health care providers at the CR center before completing CR. As a result, they did not necessarily consider the goals to be their own. Additionally, the opportunity to change goals along the way was raised as a potential improvement. Unforeseen events may occur at any time, which may affect the possibility of achievement and ownership of previously set goals. This demands greater flexibility in goal setting throughout the year. Finally, in the RCT, patients in the intervention group were asked to rate self-perceived goal achievement on a Likert scale (0-100) weekly [12,13]. All patients in this study expressed that this question was difficult to answer, and several of the patients described the scale and question as abstract.

Discussion

Principal Findings

Our findings indicate that a supervisor who possesses special characteristics is crucial to receive the full benefit of an app for increasing adherence to healthy behavior after CR. Confidence in the supervisor seems to be what enables other highlighted functions and components of the app to be perceived as motivating in relation to adherence. Other features of the app highlighted by the patients were that the app made it possible to provide individualized feedback and the use of the app was based on own goals. Additionally, the app provided a form of commitment, which proved to be of importance. Finally, to succeed in the hard work of adherence to healthy behavior after CR, patients highlighted the importance of gradually phasing out the follow-up and feedback from the supervisor.

All patients in this study highlighted the importance of the person behind the app. They described how their experiences with the supervisor's engagement, care, and support, as well as professional competence promoted motivation to adhere to healthy behavior. The trust-based relationship between the patient and supervisor could be considered a prerequisite for other components of the intervention promoting motivation to adhere to healthy behavior. To our knowledge, no qualitative studies evaluating patients' experiences using apps have clearly stated the essence of the person behind the technology. On the other hand, this finding is not surprising, as a concept analysis of adherence in the context of cardiovascular risk reduction states that adherence implies active participation and collaboration and is dependent on a concordant relationship between the patient and the health care provider [18]. A trustful relationship with a health care provider has been considered crucial for establishing strong adherence to healthy behavior [18]. An ongoing collaborative relationship between the patient and the health care provider is considered one of the most important attributes of successful adherence [19].

In this study, all patients experienced the feedback as particularly meaningful because it was individually tailored. Individual tailoring demands a person behind the app who administers the feedback. Feedback has been emphasized in the framework for the development of mobile technology use in CR [5]. In particular, individualized feedback has been proposed as a superior technique for long-term success [5,20], and may reflect the attributes of ongoing support and collaboration with a health care provider [18]. It may also reflect the supervisor's ability to influence the patients' self-efficacy [21]. People with high self-efficacy are more likely to believe that they can change their behavior than people with low self-efficacy. A positive association between self-efficacy and adherence to exercise has been described in people with coronary heart disease [22]. This is in line with a narrative review that states the importance of self-efficacy in exercise adherence among patients with chronic heart failure [23]. Exercise and exercise-based CR, which improve physical function, seem to be beneficial in order to increase self-efficacy in exercise adherence [23]. We also believe that a prerequisite for the supervisor to succeed in strengthening the patient's self-efficacy using an app is patient participation in an exercise-based CR program prior to the follow-up with the app, as in our study. Patients in our RCT were recruited from exercise-based CR programs. One of the centers documented significant improvement in peak oxygen uptake after a 12-week CR program [24], which is likely a great booster of self-efficacy.

Another factor mentioned by most patients as important for generating motivation was that the app and the follow-up

provided a commitment. The commitment was 3 fold, where the commitment to the person behind the app seemed to be the strongest. However, commitment to oneself also evolved as an important factor. The possibility of the app to aid in self-monitoring worked as a personal challenge and was described to be of value to adherence. We believe that this finding can be understood in the light of the app providing internal motivation. Internally motivated changes are considered significant for success in adherence to long-term behavioral changes [18,25].

Another attribute of successful adherence is experiencing the achievements of one's goals [19]. Most of the patients mentioned that it was important that the app and the follow-up were based on their own goals. Some even felt that this was essential to promote motivation. Goal setting is established as an effective technique in behavioral change, and setting specific goals has been shown to be effective for increasing patients' levels of physical activity after CR in terms of both frequency and duration [20]. However, guidance in setting goals that are small, important for the patient, specific, and achievable is essential to succeed with the technique [26]. Even though both CR centers included in this study considered goal setting with the patient important and the supervisor was an experienced physiotherapist from CR, some patients still mentioned an absence of ownership to their goals. Goal setting seems to be of great importance, and strategies for the implementation of the process should be highlighted in future similar interventions in the field of adherence. The importance of ownership to one's goals should not be underestimated. To maintain goals as a motivating factor for adherence to long-lasting interventions, there is a need for flexibility in terms of changing goals in line with changing needs.

The use of behavior change theory in crafting interventions has shown more powerful effects compared with interventions not based on theory [27]. The same applies to technology-based interventions. Applying behavior change theory is associated with an increased likelihood of effects in technology-based interventions [28]. The theoretical framework is important in understanding how changes are achieved [28,29]. The intervention evaluated in this study was based on the transtheoretical model (TTM) of behavior change, also known as the stages of change model [30]. According to this model, behavioral change is a process that rarely occurs in a linear manner [30]. Some of the patients experienced unforeseen events resulting in setbacks during the process toward permanent changes. They described the app and the tailored follow-up in the setback stage as a lifebuoy that helped them come back on the right track. The TTM emphasizes that setbacks in terms of moving back to a lower stage of change, that is, from the stage of maintenance to the stage of action or preparation, are more common than unusual [30]. Further, the TTM emphasizes that the need for support may be different at different stages and should be tailored to increase the likelihood of successful behavior change [30].

Interestingly, no patients suggested technical improvements of the app directly. However, many patients mentioned that the weekly rating (0-100) of self-perceived goal achievement was difficult and pointless. Therefore, a concrete improvement of the intervention would be the removal of this component. Overall, the satisfaction with the intervention, including the technical solution of the app, was high, and the use of the app was high [12]. We believe that a reason for this was that the RCT followed the Medical Research Council complex intervention framework [29,31], that is, careful and structured development of the intervention based on an evidence base and a theoretical framework [11,32,33]. A greater degree of ownership of goals was another suggestion for improvement. This will be carefully assessed and taken into account in our future planned implementation study. Additionally, we believe that a potential improvement could be the assignment of the patient's supervisor based on the patient's goals. For example, it can be beneficial if the supervisor is a nutritionist when the patient's goals are primarily diet related. This was not explicitly mentioned by the patients, but is based on the fact that more than half of the patients had goals related to weight loss, and results from the RCT did not demonstrate any statistically significant effect on bodyweight [12].

It is difficult to state whether our findings are unique as few comparable studies exist. However, a recently published systematic qualitative grounded theory review aimed at investigating the barriers to and facilitators of technology in CR and self-management [34] supports our findings. Background knowledge, ongoing support, and in-the-moment understanding, as well as personalization and gamification were concluded as facilitators [34].

Methodological Reflections and Limitations

The strength of this study is that patients from all 3 CR programs were invited to participate in the interviews, which represented the heterogeneity of patients in CR (both genders, younger and older patients, and patients living in rural and urban areas). This strengthens the credibility of the data. However, few women and few patients who originally attended the 1-week CR program were included. Their experiences are therefore represented to a lesser extent when compared with that for men and patients who originally attended the 4-week or 12-week CR program. The project leader (PL) strived to recruit more women and more patients originally from the 1-week CR program, but due to the inclusion criterion of living nearby Oslo, it was not possible. To ensure trustworthiness, all authors collaborated on the data analysis. The fact that 5 researchers conducted the analysis is expected to strengthen the dependability and overall trustworthiness. The sample size can be regarded as small, but the interviews were nuanced, and we considered the material to be saturated after 8 interviews. This view was also valid after 10 interviews when we decided to end the data collection.

The purpose of qualitative research is directed toward providing in-depth explanations and meanings rather than generalizing findings [35]. The term "transferability" is used to express to what degree the findings can be applied to other contexts. The transferability of this study has to be judged by the reader. We hope to have highlighted some phenomena that may have relevance for comparable patient populations and situations, such as app-based interventions aiming to promote adherence to healthy behavior in patients with lifestyle diseases.

Since the interviews were conducted after the end of the intervention in the RCT, oversights and recall biases of relevant experiences and suggestions for improvement cannot be ruled out. Interviews during different phases of the intervention (ie, after 3, 6, and 12 months) could have resulted in more accurate snapshots of the patients' experiences.

Regarding the positions and preconceptions of the researchers, the first author's first-hand experiences with the intervention through being the project coordinator and supervisor for all patients included in the RCT may have had an influence. For example, the overall idea of evaluating patients' experiences and thereby the choice of the research question was based on regular feedback from patients during the RCT. Further, the engagement of 2 authors (AB and BBN) in the RCT, which this study builds on, may also have had an influence. Even though all the authors have professional and research interests in the field of health science, there was diversity among author backgrounds (physiotherapy and dietetics), as well as diversity in relation to author experiences with the use of technology and their previous engagement in the RCT. This led to interesting discussions and enhanced reflexivity [36]. The overall experience of the researchers of this study most likely indicates that there were certain things that we took for granted. However, it also means that we were well positioned to understand the context and to perform the study [37].

Conclusions and Implications

Overall, appreciation of the person behind the app turned out to be a consistent finding. This person seems to promote motivation in general and seems to enable other known behavioral change techniques to be motivating, such as feedback and goal setting. Therefore, the person behind the app (the supervisor) seems to be one of the main reasons for the high acceptance and use of the app, and consequently, is important for the results in the RCT. We therefore conclude that health care providers should actively participate in the patients' process of adherence and that the use of the app should not be considered a substitute but a reinforcement in motivational work to promote adherence to healthy behavior after CR. Future development of interventions in the field of adherence should therefore strive to create tools that enable an ongoing collaborative relationship between the patient and the health care provider; provide follow-up based on patients' own goals, of which they have ownership; and provide feedback and support to patients at the stage of change, at any given time.

Acknowledgments

The authors would like to express their gratitude to the patients who participated and made this study possible. We confirm that all personal identifiers have been removed or disguised so the patients described are not identifiable and cannot be identified through the presented details. This project has been made possible by Dam Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Thematic interview guide.
[DOCX File , 18 KB - humanfactors_v9i1e34294_app1.docx]

References

1. Ambrosetti M, Abreu A, Corrà U, Davos CH, Hansen D, Frederix I, et al. Secondary prevention through comprehensive cardiovascular rehabilitation: From knowledge to implementation. 2020 update. A position paper from the Secondary Prevention and Rehabilitation Section of the European Association of Preventive Cardiology. Eur J Prev Cardiol 2020 Apr 07:181-205. [doi: 10.1177/2047487320913379] [Medline: 33611446]
2. Authors/Task Force Members:, Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, et al. Atherosclerosis 2016 Sep;252:207-274. [doi: 10.1016/j.atherosclerosis.2016.05.037] [Medline: 27664503]
3. Piepoli MF, Corrà U, Adamopoulos S, Benzer W, Bjarnason-Wehrens B, Cupples M, et al. Eur J Prev Cardiol 2014 Jun;21(6):664-681. [doi: 10.1177/2047487312449597] [Medline: 22718797]
4. Kotseva K, Wood D, De Bacquer D, De Backer G, Rydén L, Jennings C, EUROASPIRE Investigators. EUROASPIRE IV: A European Society of Cardiology survey on the lifestyle, risk factor and therapeutic management of coronary patients from 24 European countries. Eur J Prev Cardiol 2016 Apr;23(6):636-648. [doi: 10.1177/2047487315569401] [Medline: 25687109]
5. Beatty AL, Fukuoka Y, Whooley MA. Using mobile technology for cardiac rehabilitation: a review and framework for development and evaluation. J Am Heart Assoc 2013 Nov 01;2(6):e000568 [FREE Full text] [doi: 10.1161/JAHA.113.000568] [Medline: 24185949]
6. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]

7. Gandhi S, Chen S, Hong L, Sun K, Gong E, Li C, et al. Effect of Mobile Health Interventions on the Secondary Prevention of Cardiovascular Disease: Systematic Review and Meta-analysis. Can J Cardiol 2017 Feb;33(2):219-231. [doi: 10.1016/j.cjca.2016.08.017] [Medline: 27956043]
8. Marzano L, Bardill A, Fields B, Herd K, Veale D, Grey N, et al. The application of mHealth to mental health: opportunities and challenges. The Lancet Psychiatry 2015 Oct;2(10):942-948. [doi: 10.1016/S2215-0366(15)00268-0] [Medline: 26462228]
9. Woods LS, Duff J, Roehrer E, Walker K, Cummings E. Patients' Experiences of Using a Consumer mHealth App for Self-Management of Heart Failure: Mixed-Methods Study. JMIR Hum Factors 2019 May 02;6(2):e13009 [FREE Full text] [doi: 10.2196/13009] [Medline: 31045504]
10. Georgsson M, Staggers N. Patients' Perceptions and Experiences of a mHealth Diabetes Self-management System. Comput Inform Nurs 2017 Mar;35(3):122-130. [doi: 10.1097/CIN.0000000000000296] [Medline: 27748662]
11. Lunde P, Nilsson BB, Bergland A, Bye A. Feasibility of a Mobile Phone App to Promote Adherence to a Heart-Healthy Lifestyle: Single-Arm Study. JMIR Form Res 2019 Apr 19;3(2):e12679 [FREE Full text] [doi: 10.2196/12679] [Medline: 31066684]
12. Lunde P, Bye A, Bergland A, Grimsmo J, Jarstad E, Nilsson BB. Long-term follow-up with a smartphone application improves exercise capacity post cardiac rehabilitation: A randomized controlled trial. Eur J Prev Cardiol 2020 Nov;27(16):1782-1792 [FREE Full text] [doi: 10.1177/2047487320905717] [Medline: 32106713]
13. Lunde P, Bye A, Bergland A, Nilsson BB. Effects of individualized follow-up with a smartphone-application after cardiac rehabilitation: protocol of a randomized controlled trial. BMC Sports Sci Med Rehabil 2019;11:34 [FREE Full text] [doi: 10.1186/s13102-019-0148-2] [Medline: 31768261]
14. Park LG, Beatty A, Stafford Z, Whooley MA. Mobile Phone Interventions for the Secondary Prevention of Cardiovascular Disease. Prog Cardiovasc Dis 2016;58(6):639-650 [FREE Full text] [doi: 10.1016/j.pcad.2016.03.002] [Medline: 27001245]
15. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
16. NVivo. QSR International. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2022-03-14]
17. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. Cohen SM. Concept analysis of adherence in the context of cardiovascular risk reduction. Nurs Forum 2009;44(1):25-36. [doi: 10.1111/j.1744-6198.2009.00124.x] [Medline: 19187051]
19. Bosworth H, Oddone E, Weinberger M. Patient Treatment Adherence: Concepts, Interventions, and Measurement. Hove, East Sussex, UK: Psychology Press; 2006.
20. Ferrier S, Blanchard CM, Vallis M, Giacomantonio N. Behavioural interventions to increase the physical activity of cardiac patients: a review. Eur J Cardiovasc Prev Rehabil 2011 Feb;18(1):15-32. [doi: 10.1097/HJR.0b013e32833ace0e] [Medline: 20502341]
21. Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review 1977;84(2):191-215. [doi: 10.1037/0033-295x.84.2.191]
22. Slovinec D'Angelo ME, Pelletier LG, Reid RD, Huta V. The roles of self-efficacy and motivation in the prediction of short- and long-term adherence to exercise among patients with coronary heart disease. Health Psychol 2014 Nov;33(11):1344-1353. [doi: 10.1037/hea0000094] [Medline: 25133848]
23. Ha FJ, Hare DL, Cameron JD, Toukhsati SR. Heart Failure and Exercise: A Narrative Review of the Role of Self-Efficacy. Heart Lung Circ 2018 Jan;27(1):22-27. [doi: 10.1016/j.hlc.2017.08.012] [Medline: 28969981]
24. Nilsson BB, Lunde P, Holm I. Implementation and evaluation of the Norwegian Ullevaal model as a cardiac rehabilitation model in primary care. Disabil Rehabil 2019 Feb;41(4):481-488. [doi: 10.1080/09638288.2017.1397776] [Medline: 29105515]
25. Ryan RM, Lynch MF, Vansteenkiste M, Deci EL. Motivation and Autonomy in Counseling, Psychotherapy, and Behavior Change: A Look at Theory and Practice 1ψ7. The Counseling Psychologist 2010 Feb 10;39(2):193-260. [doi: 10.1177/0011000009359313]
26. Miller W, Rollnick S. Motivational Interviewing: Helping People Change. New York, NY: Guilford Press; 2013.
27. Glanz K, Rimer B, Viswanath K. Health Behavior: Theory, Research, and Practice. Hoboken, NJ: John Wiley & Sons; 2015.
28. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
29. 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]
30. Prochaska J, Redding C, Evers K. The Transtheoretical Model and Stages of Change. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior: Theory, Research, and Practice. Hoboken, NJ: Wiley; 2015:125-148.

31. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
32. Richards D, Hallberg I. Complex Interventions in Health: An overview of research methods. Oxfordshire, England, UK: Routledge; 2015.
33. Lunde P, Nilsson BB, Bergland A, Kværner KJ, Bye A. The Effectiveness of Smartphone Apps for Lifestyle Improvement in Noncommunicable Diseases: Systematic Review and Meta-Analyses. J Med Internet Res 2018 May 04;20(5):e162 [FREE Full text] [doi: 10.2196/jmir.9751] [Medline: 29728346]
34. Tadas S, Coyle D. Barriers to and Facilitators of Technology in Cardiac Rehabilitation and Self-Management: Systematic Qualitative Grounded Theory Review. J Med Internet Res 2020 Nov 11;22(11):e18025 [FREE Full text] [doi: 10.2196/18025] [Medline: 33174847]
35. Carminati L. Generalizability in Qualitative Research: A Tale of Two Traditions. Qual Health Res 2018 Nov;28(13):2094-2101. [doi: 10.1177/1049732318788379] [Medline: 30043686]
36. Malterud K. Qualitative research: standards, challenges, and guidelines. The Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]
37. Berger R. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. Qualitative Research 2013 Jan 03;15(2):219-234. [doi: 10.1177/1468794112468475]

Abbreviations

CR: cardiac rehabilitation
mHealth: mobile health
OsloMet: Oslo Metropolitan University
RCT: randomized controlled trial
TTM: transtheoretical model

Edited by A Kushniruk; submitted 15.10.21; peer-reviewed by V Buss, F Velayati; comments to author 07.12.21; revised version received 28.01.22; accepted 03.02.22; published 23.03.22.

Please cite as:
Lunde P, Bye A, Bruusgaard KA, Hellem E, Nilsson BB
Patients' Experiences of Using a Smartphone App After Cardiac Rehabilitation: Qualitative Study
JMIR Hum Factors 2022;9(1):e34294
URL: https://humanfactors.jmir.org/2022/1/e34294
doi:10.2196/34294
PMID:35319476

Original Paper

Designing Formulae for Ranking Search Results: Mixed Methods Evaluation Study

Laura Douze[1,2*], MSc; Sylvia Pelayo[1,2*], PhD; Nassir Messaadi[3*], MD, PhD; Julien Grosjean[4,5*], PhD; Gaétan Kerdelhué[4,5*], BSc; Romaric Marcilly[1,2*], PhD

[1]Inserm, Centre d'Investigation Clinique pour les Innovations Technologiques 1403, Institut Coeur-Poumon, Lille, France

[2]Unité Labellisée de Recherche 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, Univ. Lille, Centre Hospitalier Universitaire de Lille, Lille, France

[3]Département de médecine générale, Univ. Lille, Lille, France

[4]Département d'Informatique et d'Information Médicales, Centre hospitalier universitaire de Rouen, Rouen, France

[5]Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, Inserm 1142, Sorbonne Paris Nord University, Sorbonne Paris Cité, Villetaneuse, France

[*]all authors contributed equally

Corresponding Author:
Laura Douze, MSc
Inserm, Centre d'Investigation Clinique pour les Innovations Technologiques 1403
Institut Coeur-Poumon
3ème étage Aile Est, CS 70001
Bd du Professeur Jules Leclercq
Lille, 59037
France
Phone: 33 0362943939
Email: laura.douze@univ-lille.fr

Abstract

Background: A major factor in the success of any search engine is the relevance of the search results; a tool should sort the search results to present the most relevant documents first. Assessing the performance of the ranking formula is an important part of search engine evaluation. However, the methods currently used to evaluate ranking formulae mainly collect quantitative data and do not gather qualitative data, which help to understand what needs to be improved to tailor the formulae to their end users.

Objective: This study aims to evaluate 2 different parameter settings of the ranking formula of LiSSa (the French acronym for *scientific literature in health care*; Department of Medical Informatics and Information), a tool that provides access to health scientific literature in French, to adapt the formula to the needs of the end users.

Methods: To collect quantitative and qualitative data, user tests were carried out with representative end users of LiSSa: 10 general practitioners and 10 registrars. Participants first assessed the relevance of the search results and then rated the ranking criteria used in the 2 formulae. Verbalizations were analyzed to characterize each criterion.

Results: A formula that prioritized articles representing a consensus in the field was preferred. When users assess an article's relevance, they judge its topic, methods, and value in clinical practice.

Conclusions: Following the evaluation, several improvements were implemented to give more weight to articles that match the search topic and to downgrade articles that have less informative or scientific value for the reader. Applying a qualitative methodology generates valuable user inputs to improve the ranking formula and move toward a highly usable search engine.

(JMIR Hum Factors 2022;9(1):e30258) doi:10.2196/30258

KEYWORDS

information retrieval; search engine; topical relevance; search result ranking; user testing; human factors

Introduction

Background

The evolution of the World Wide Web from a static network (Web 1.0) to a semantic web (Web 3.0) is ever more palpable [1]. The semantic web provides access to information in billions of heterogeneous documents in various formats, stored on different operating systems, and references among others to varying extents. This opens up a range of possibilities such as facilitating rapid access to targeted data [1]. However, the challenge for health care professionals is to identify relevant documents in this ocean of data [2,3].

In this context, search engine evaluation and improvement are key issues [4]. As soon as the discipline of information retrieval was established, researchers started to combine structured evaluation methods. For example, the Cranfield method (developed in 1962 [5]) soon became a benchmark for evaluating information retrieval and the Text Retrieval Conference has encouraged initiatives in information retrieval since 1992 [6].

Since then, the methods for evaluating information retrieval have diversified to meet a broader range of objectives. There are two main types of evaluation: system-oriented evaluations [4,7] that focus on search engine optimization (search efficiency, recall, accuracy, etc) and user-oriented evaluations [4,7,8] that seek to improve the user experience and search engine's value (usability, expressivity, relevance, etc). One of the most important factors, perhaps the most important factor for search engines, is the relevance of search results [9].

There are two main definitions of relevance [10]: objective relevance (ie, the search result contains the submitted keyword) and subjective relevance (ie, the search result satisfies the user). Subjective relevance can then be subdivided into four main categories [10]: topical, situational, motivational, and affective. Topical relevance is the most studied type [4,8] and is the subject of this study; it was defined by Harter [11] as "how well the topic of the information retrieved matches the topic of the request."

It is possible to evaluate topical relevance by involving users (eg, when relevance is rated by one or more expert or nonexpert participants) [12,13] or without their involvement (eg, in batch evaluations, such as the Cranfield method). Conventional methods for evaluating the relevance or performance of search engines are mostly based on comparisons between several formulae or a comparison with a gold standard. These comparisons are performed with quantitative data (mostly judges' ratings) [4,14,15]. This method generates a large amount of data. The evaluation is quick and can be performed remotely. Thus, it is possible to include a large number of judges and test a large number of search queries. However, this method does not provide qualitative data, information on why a formula fails, or information on how to improve a formula's performance.

Some studies have included user feedback, that is, the collection of qualitative data on perceived relevance and judgment criteria [16,17]. However, to the best of our knowledge, most of these studies sought to model and understand users' relevance judgments rather than to evaluate and improve existing ranking formulae. This is a shortcoming of current methods for improving sorting formulae. Qualitative methods should also be used to identify the strengths and weaknesses of formulae.

In human factors research, it is well known that participative methods (notably user-centered designs involving users at each step in the design process [18]) improve the usability of a product before implementation in real settings. If users are not involved in the design process, their needs are often hypothetical and come from designers' own representations of the field [19]. The tools thus created may not correspond to the users' true needs and habits, which typically creates usability problems. Iterative evaluations are needed to improve effectiveness, efficiency, utility, acceptability, end user satisfaction, and (in health care) the safety of health care professionals and patients [20-22]. A proven method is user testing (also known as usability testing), which "calls for representative users to perform representative tasks as a means to reveal the interactive strengths and opportunities for improvement of a device" [23]. When coupled with the think-aloud method, a verbal report method from cognitive psychology that provides information on the cognitive behavior of participants performing a task [24], user testing collects valuable qualitative data about users' behaviors and needs. Given that the user and moderator can interact during the evaluation, the user's behavior and verbalizations can be investigated directly and may help clarify the user's responses.

With a view to prompting further design innovations, we describe here the formative assessment of the *sort by relevance* function of a health care literature search engine. Taking a broader view, we developed a 2-step methodology that lies between a conventional information retrieval approach (for evaluating the relevance of search results) and a conventional human factors approach (for evaluating the usability of a new technology). With the objective of improving the ranking and moving toward a useful, usable search interface, we collected data on the performance of 2 *sort by relevance* formulae and on their strengths and weaknesses. We focused on the value of active end user involvement in this evaluation as a means of improving topical relevance in comparison with common evaluation methods used in the field.

Study Context

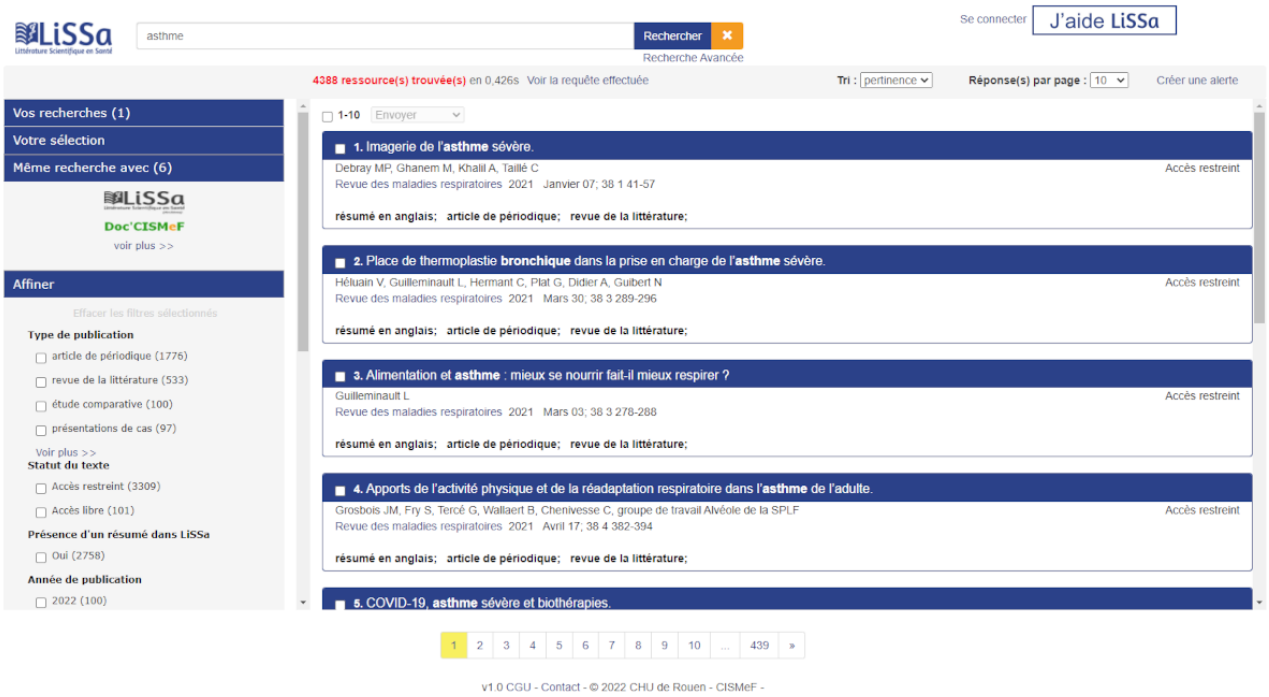
This study was part of a broader research program funded by the French National Research Agency. The objective of the project is to develop a health care literature search engine *LiSSa*, the French acronym for *scientific literature in health care* (LiSSa.fr [25]). The particularity of the search engine is that both the interface and content are entirely in French.

PubMed is the most widely used search engine for scientific literature on health care. It is an important tool for all health professionals, for lifelong training and for updating their knowledge. However, English is a hindrance to reading by many French health care professionals [26]. French professionals are often not sufficiently fluent in English to read scientific articles. For these professionals, the lack of tools in French that allows them to find scientific literature in their native language is an obstacle to updating their medical knowledge and continuing education [26]. LiSSa is a French language tool that provides

access to French scientific literature on health to people who are not specialized in scientific research and who do not understand English well enough. The main target users are general practitioners (GPs) and hospital registrars for continuing education, updating knowledge, and helping them find scientific articles to solve medical issues. In short, the tool helps them in the context of daily practice outside any academic or institutional environment. LiSSa currently encompasses over 1,300,000 French scientific references provided by various publishers and sources, among which the PubMed database (US National Library of Medicine [NLM]) accounts for 53% and publisher Elsevier accounts for 23% (18% without overlap).

The project is led by the Department of Medical Informatics and Information (D2IM). This department from the Rouen University hospital specializes in eHealth, more precisely in knowledge representation (terminologies, ontologies, etc) and information management (databases and search engines). The D2IM design team comprises physicians, librarians, and computer scientists. Their previous work includes Catalogue and Index of French Language Health Resources on the Internet and the Health Terminology and Ontology Portal [27,28], both available on the web. D2IM created the LiSSa database and designed the graphical user interface for the tool. The other academic partner is the Clinical Investigation Centre for Innovative Technology (for *Clinical Investigation Centre for Innovative Technology* in French) of the Lille University Hospital, an academic research laboratory that works to improve the design and evaluation of innovations in health care and is responsible for the usability assessments of the tool. In total, 3 companies were partners in the project: Elsevier Masson, one of the world's leading science publishers, and the French start-ups Alicante and Sensegate. The LiSSa.fr [25] website (Figure 1) was published in 2014 [29]. Initial evaluations by GPs revealed a lack of relevance to the search results. They considered that the sorted results did not present the most relevant articles first; the top-ranked articles were often of little practical value, too old, or not representative of the topic. Specific work to improve the sorting of results needs to be conducted.

Figure 1. A screenshot of a search results page of LiSSa.fr [25].



The most commonly used search engines use several criteria to rank database search results: the match between the keywords used and the metadata, the number of views, and contextual data (such as the user's previous search or geographic location) [30-32]. However, the best results are typically obtained using a combination of ranking criteria. With regard to the LiSSa search engine, D2IM considered two ranking formulae, A and B, which differed in the weight attributed to the same set of criteria:

- Formula A prioritizes recent novel articles by assigning more weight to the publication year.
- Formula B prioritizes general articles that represent consensus in the field (literature reviews, meta-analyses, etc) by assigning more weight to the publication type.

The formulae's weighting criteria are listed in Table 1. All these criteria are based on metadata retrieved from publishers and PubMed (produced by the NLM) [26]. For papers not indexed in PubMed, metadata were automatically generated with a set of indexing terms from the NLM's controlled vocabulary, used to index articles in the biomedical field (MeSH [Medical Subject Headings] thesaurus) [33].

Table 1. Weighting of each criterion in the ranking formulae A and B.

Criterion	Weighting for formula A	Weighting for formula B
Title	10	10
Subtitle	10	10
Author keywords	5	5
Major MeSH[a] terms[b]	4	4
Minor MeSH terms	1	1
Nonexploded indexing	3	3
Exploded indexing[c]	1	1
Manual indexing[d]	3	3
Automatic indexing	1	1
Year of publication	10 for the current year and −2 for each year in the past	10 for the current year and −0.6 for each year in the past
Type of publication; for example, good practice guidelines, consensus statements, directives, literature reviews, and meta-analyses	0	3

[a]MeSH: Medical Subject Headings.

[b]In the field of biomedicine, articles are often indexed according to the MeSH thesaurus. LiSSa considers the MeSH terms to be major when they correspond to one of the article's main themes or minor when they correspond to one of the article's subthemes.

[c]The MeSH thesaurus is structured like a tree; an MeSH term typically has several hierarchical levels above and below it. For example, *asthma* belongs to the *bronchial diseases* category and one of its narrower terms is *status asthmaticus*. A search for asthma will thus also find an article indexed as *status asthmaticus* but the latter will be less weighted because indexing is said to be exploded.

[d]Some documents are indexed by a National Library of Medicine indexer; this is referred to as manual indexing. Other documents are indexed by text mining tools, which is referred to as automatic indexing. Manual indexing is considered to be more accurate and efficient than automatic indexing.

Objectives

The goals of this study are to (1) determine which formula (A or B) is associated with the greatest topical relevance and (2) adjust the ranking formula's criteria to meet the target end users' needs more closely.

Methods

Overview

To evaluate the ranking formulae, we conducted formative user testing, which consisted of directly observing users using the tool in a controlled situation. It is a well-known method used in human factors to collect user behaviors and identify their needs [24].

Here, the user tests were conducted in two steps:

1. User evaluations of the two ranking formulae: the participants had to rate the relevance of search results produced by the two ranking formulae, while justifying their ratings. This step enabled us to determine which formula was associated with the greatest topical relevance.
2. Data collection for improving the ranking formulae: the participants had to rate the ranking criteria used in both formulae A and B and some additional criteria in terms of establishing an article's relevance (from the most important criterion to the least important). Coupled with the users' verbalizations when rating relevance, these data enabled us to adjust the ranking formula's criteria and thus develop a formula that, in principle, would match professionals' needs more closely.

Step 1: Comparison of the 2 Ranking Formulae

Data Collection

LiSSa.fr [25] search logs were searched to identify the most frequent queries made by users. Among them, 2 were selected by a GP from the project consortium for their potential clinical value for the participants. Half of the participants (10/20, 50%) used the search query 1 (*treatment-resistant depression*) and the other half used search query 2 (*sleep apnea syndrome*; Table 2).

To compare the performances of ranking formulae A and B, each participant successively evaluated the search results generated by formulae A and B for the same query. The order of presentation of the formulae was counterbalanced to avoid order effects (Table 2). Participants were asked to perform their search query with LiSSa and then rate the relevance of the first 10 search results on a 5-point Likert scale ranging from 1 (not at all relevant) to 5 (highly relevant). For each result, the users had to justify their choice (eg, the article was too old or off-topic). The verbalizations for each rating were recorded.

After using both formulae, the users were asked to rate their overall level of satisfaction on a 7-point Likert scale ranging from 1 ("I am not at all satisfied with the search results") to 7 ("I am fully satisfied with the search results"). For greater discriminative power, we chose to use a 7-point scale.

Table 2. Distribution of the participants according to the order in which the formulae and the predetermined search queries were presented (N=20).

Order of tested formula and predetermined search query	General physician participants, n (%)	Registrar participants, n (%)
Tested formula A, then B		
Query 1	3 (15)	2 (10)
Query 2	2 (10)	3 (15)
Tested formula B, then A		
Query 1	2 (10)	3 (15)
Query 2	3 (15)	2 (10)

Data Analysis

To check whether the *query* (query 1 vs query 2) and *type of participant* (GPs vs registrars) variables did not have an effect, a Mann–Whitney *U* test for independent samples was performed on the difference in scores between the formulae A and B.

To compare the user-perceived relevance for each formula, a Mann–Whitney *U* test for matched samples was performed on the three data sets:

- The scores given to the first 10 articles.
- The normalized discounted cumulative gain (NDCG) [34] was calculated from article scores. NDCG is an equation that calculates a score between 0 and 1; it evaluates the relevance of the article ranking using the scores given by the participants. Hence, the NDCG is close to 1 when the highest-rated articles are presented before the lowest-rated articles and, on the contrary, is close to 0 when the formula presents low-rated articles first and high-rated articles last.
- The overall satisfaction score awarded by the user at the end of the testing.

All statistical analyses were performed using R software (R Foundation for Statistical Computing) [35]. The threshold for statistical significance was set at $P<.05$ in all tests.

Participants' verbalizations when justifying their scores were thematically analyzed [36]. This analysis enabled us to identify the strengths and weaknesses of each ranking formula, based on positive or negative comments. Each theme was counted once for each participant.

Step 2: Prioritization of the Ranking Criteria

Data Collection

To improve the relevance of the ranking formulae and refine the criteria and their respective weightings, the participants were shown a list of criteria on separate cards, with the name of the criterion on one side and its explanation (that the user could consult, if required) on the other side (Table 3). The criteria were presented in random order for each participant. Additional explanations were provided upon request. The list contained both the criteria already included in formulae A and B and several other potentially relevant criteria.

The users were asked to classify the criteria by order of importance and to justify their choices. The justifications for each criterion were noted by specifying the item's valence (positive or negative comments).

Table 3. List of the criteria shown to the participants.

Name	Explanation
Title	The keyword is present in the article's title.
Subtitle	The keyword is present in the article's subtitle.
Author keywords	The keyword is present in the author keywords.
Abstract	The keyword is present in the article's abstract.
Major MeSH[a] term	The keyword is present in the major MeSH term.
Minor MeSH term	The keyword is present in the minor MeSH term.
Exploded indexing or not[b]	Points are awarded if the indexing is not exploded (the keyword is the same as the MeSH term) vs exploded indexing (the keyword is found among the narrower MeSH terms).
Manual or automatic indexing	Points are awarded if the indexing is manual (performed by a National Library of Medicine indexer) rather than automatic (performed by text mining).
Association with a qualifier	Points are subtracted if the indexing qualifier is specified: for example, with *asthma/diagnosis*, the article will deal only with the diagnosis of asthma and not with asthma in general.
Year of publication	Points are awarded as a function of the article's year of publication: the more recent it is, the more points it will be awarded.
Type of publication	Points are awarded if the article is a literature review, a good practice guideline, a consensus statement, a directive, or a meta-analysis.
Presence of an abstract	Points are awarded if an abstract in French is directly available on LiSSa (ie, without having to visit the journal's website).
The journal's importance	Points are awarded as a function of the journal's impact.

[a]MeSH: Medical Subject Headings.

[b]The MeSH thesaurus contains qualifiers that can be linked to each keyword to make it more *precise*. For example, the index entry *asthma* can be specified by the qualifier *diagnosis* (*asthma/diagnosis*), to tell the reader that only the diagnosis of asthma is addressed in the article, and not its other aspects (treatment, complication, etc).

Data Analysis

We analyzed the classification of the ranking criteria by calculating the mean and median ranks for each criterion. Kendall *W* was used to evaluate the degree of interrater agreement.

Participants' verbalizations were analyzed to characterize each criterion's positive qualities (ie, why the user wanted to include it in the ranking formula) or negative qualities (ie, why it should not be taken into account or only partly in the formula).

Test Participants

Calls for participation were made by the Department of General Practice and Family Medicine of the University of Lille by email to recruit GPs. Announcements were made during registrar classes, and calls for participation were posted in discussion groups and on social media pages to recruit hospital registrars. The only recruitment criterion was the profile of the participant (GP or registrar).

A total of 10 GPs and 10 registrars (ie, LiSSa's target users) participated in the tests. They volunteered to participate, and no compensation was paid for their participation.

All sessions were filmed and subsequently analyzed offline by a usability engineer. The participants accessed LiSSa via a computer with an internet connection.

Ethics Consideration

This study is a human and social science study. The French law governing 'research involving the human person' exempts human and social science studies from requiring approval from an ethics committee. Written informed consents were obtained from each participant before they took part in the study.

Results

Participant Characteristics

The characteristics of the participants are shown in Table 4.

Table 4. Participant characteristics.

Participant number	Profile	Age (years)	Number of years of practice (including internship semesters for registrar)	Self-reported frequency of use of a search engine
P1	GP[a]	29	2	Frequently
P2	GP	28	0	Frequently
P3	GP	30	2.5	Frequently
P4	GP	55	26	Frequently
P5	GP	56	29	Frequently
P6	GP	68	30	Frequently
P7	GP	53	16	Frequently
P8	GP	53	25	Not often
P9	GP	55	27	Frequently
P10	GP	33	5	Frequently
P11	Registrar	24	0.5	Never
P12	Registrar	26	0.5	Never
P13	Registrar	28	4	Never
P14	Registrar	26	1.5	Frequently
P15	Registrar	30	4	Not often
P16	Registrar	26	2	Frequently
P17	Registrar	28	1.5	Frequently
P18	Registrar	25	1.5	Not often
P19	Registrar	31	4.5	Frequently
P20	Registrar	29	5	Not often

[a]GP: general practitioner.

Step 1: Comparison of the 2 Ranking Formulae

Our statistical analysis did not show a significant effect of *query* (treatment-resistant depression vs sleep apnea syndrome; W=4935; P=.87). Similarly, *the type of participant* (GPs vs registrars) did not have a significant effect (W=5071.5; P=.86). Therefore, a single user group (all 20 participants) was considered in the subsequent statistical tests.

Statistical tests showed that formula B was preferred to formula A with regard to all 3 end points (Table 5).

The analysis of the participants' verbalizations confirmed this finding (Figure 2). Formula A attracted more negative comments: more participants expressed that the articles were not useful in practice, off-target, or too specific to a given population. Concerns about an article's recentness were rarely expressed, although 15% (3/20) of the participants thought that at least one article presented by formula B was too old (Figure 2).

Most participants (14/20, 70%) preferred formula B, notably because the articles' topics were general and did not focus on a specific population (Figure 3). However, 30% (3/10) of GPs and 10% (1/10) of registrars preferred formula A because formula B presented trivial articles that taught them nothing new.

This phase of the evaluation prompted us to conclude that formula B best met participants' expectations.

Table 5. The mean and median ranking scores, the normalized discounted cumulative gain (NDCG), and overall satisfaction scores for formulae A and B (N=20 participants).

	Formula A	Formula B	*W* value	*P* value
Main ranking score, median (IQR), out of 5	3.57 (4-2.5)	3.82 (4-3.5)	3518.5	.02
Main NDCG, median (IQR), out of 1	0.87 (0.95-0.83)	0.97 (0.99-0.94)	7	.01
Overall satisfaction score, median (IQR), out of 7	4.7 (5-4.6)	5.8 (6-5.6)	27.5	.01

Figure 2. Types of negative verbalization about the articles for formula A or formula B; the number of participants is stated.

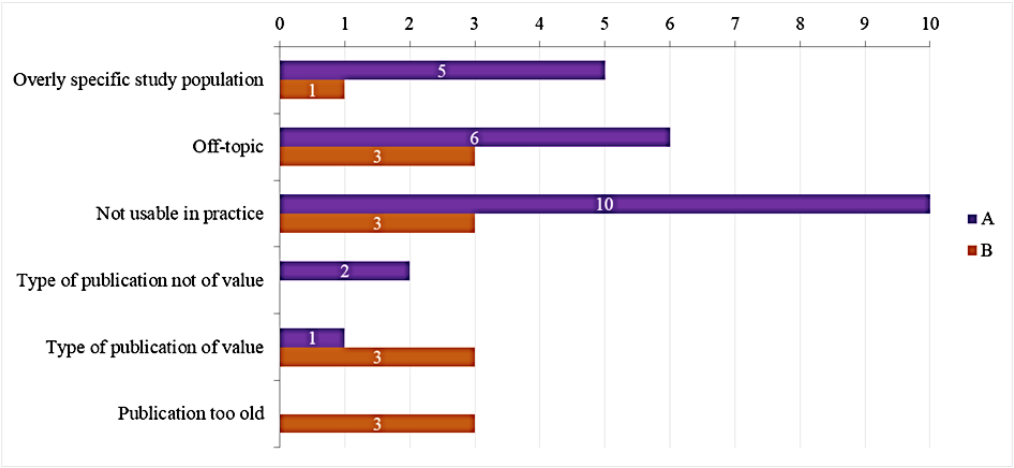
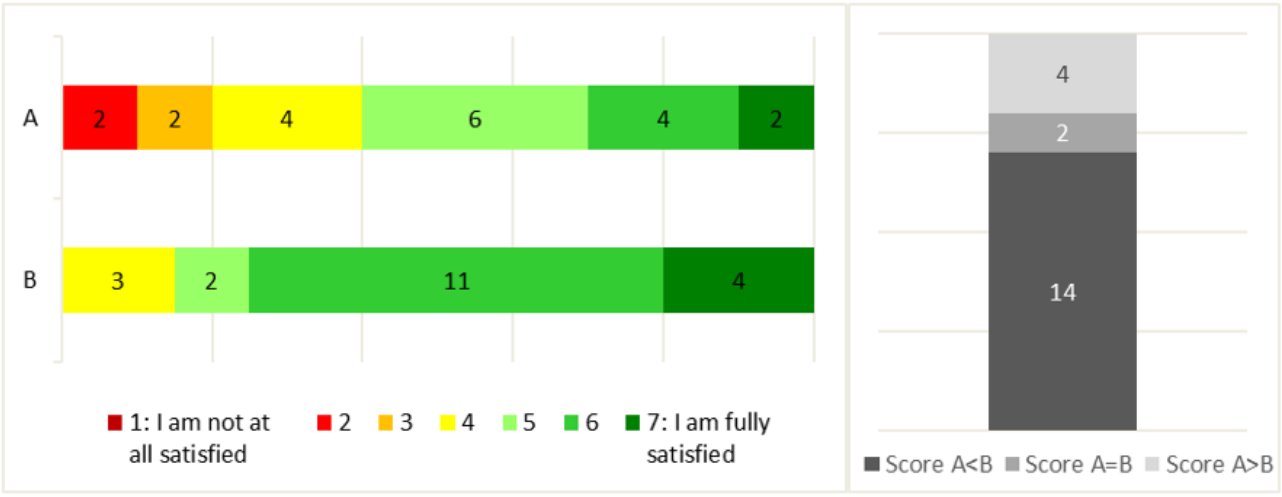


Figure 3. Distribution of the overall satisfaction score for formulae A and B (left panel), and the number of participants who gave formula A a higher, equal, or lower score than formula B (right panel).



Step 2: Prioritization of the Ranking Criteria

In step 1, dealing with the formulae and search result scores enabled users to evaluate the formulae's strengths and weaknesses. Because of the criteria ranking process, step 2 refined the users' needs by moving out of the context of the present formulae and predetermined search queries. This part of the study enabled us to adjust formula B and thus make it more closely match the users' needs. Of the 20 participants, 2 (10%) did not perform this step; hence, 18 (90%) participants (9 GPs and 9 registrars) prioritized the criteria.

The participants' mean and median criterion rankings are shown in Table 6.

Given that the absence of an effect of *participants* had not been demonstrated for this data set, GPs and registrars were considered separately. There was a low but statistically significant degree of agreement among the GPs (W=0.46; P<.001; r=.39) and registrars (W=0.35; P<.001; r=.27). Despite this low agreement, clear trends emerged in the criteria rankings (Figure 4):

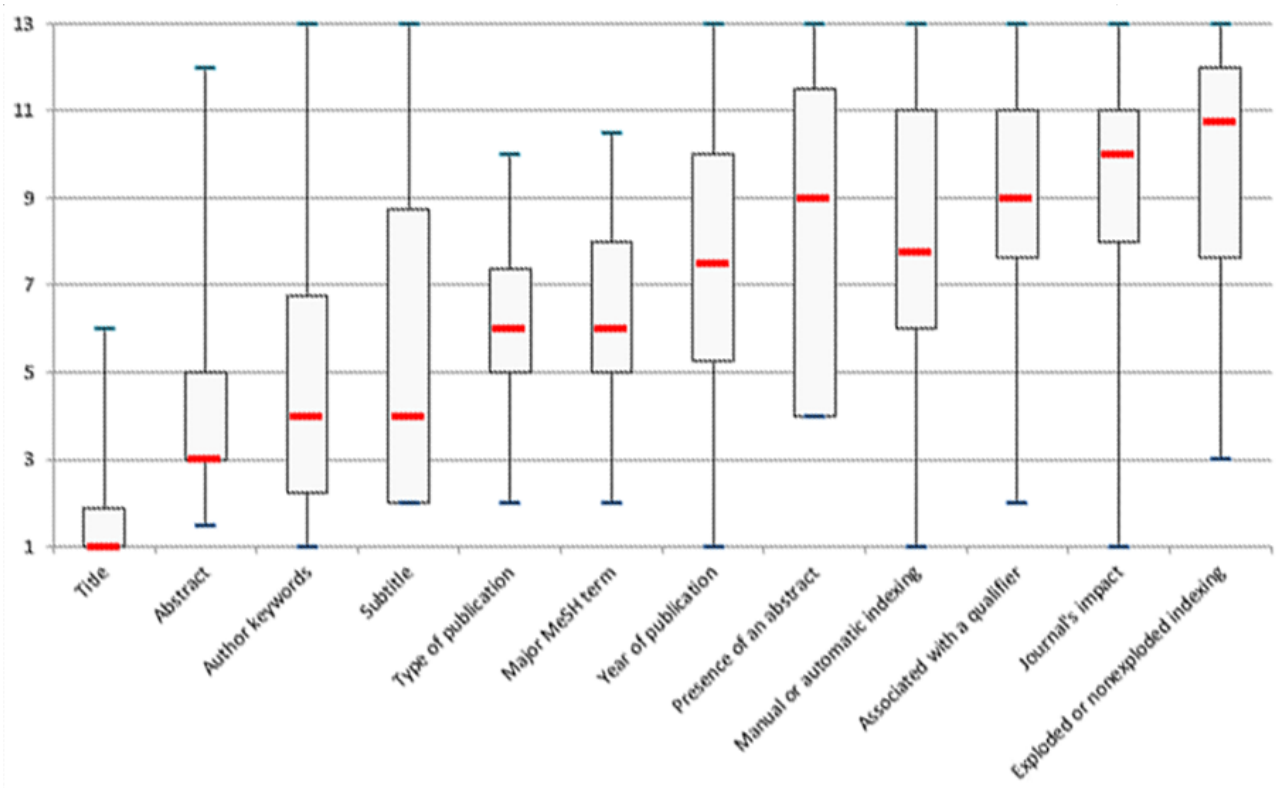
- Participants considered the *title* to be the most representative relevance criterion; it was one of the elements they looked at first, and it was thought to reflect the research rather well.
- After the title, the users looked for the presence of the query keywords in the *Abstract*, *author keywords*, and *subtitle*.
- Publication type was an important criterion because it enabled the selection of the most reliable articles (ie, those with a higher level of evidence). Participants gave low ratings to the editorials and letters.
- The *year of publication* is controversial. Some users judged it to be important because it reflected the latest advances, whereas others considered it to be highly dependent on the topic of the search query. Older articles are still used as benchmarks for practice in some fields.
- The remaining criteria were judged to be of secondary importance, albeit occasionally of value in differentiating between 2 articles with the same score. For example, articles describing more general studies were preferred to those describing more specific studies (*exploded indexing* and *associated with a qualifier*).

Table 6. Mean and median criterion ranks (n=18 participants).

Criterion	Mean rank	Median rank (IQR)
Title	1.8	1 (1.87-1)
Abstract	4.5	3 (5-3)
Author keywords	5.0	4 (6.75-4)
Subtitle	5.7	4 (6.75-2.25)
Type of publication	6.1	6 (7.38-5)
Major MeSH[a] term	6.5	6 (8-5)
Year of publication	7.8	7.5 (10-5.25)
Presence of an abstract	8.1	7.8 (11.5-4)
Manual or automatic indexing	8.1	9 (11-6)
Associated with a qualifier	8.6	9 (11-7.62)
The journal's impact	9.3	10 (11-8)
Exploded or nonexploded indexing	9.9	10 (12-7.63)
Minor MeSH term	9.9	10.8 (11.75-9)

[a]MeSH: Medical Subject Headings.

Figure 4. Boxplots of the scores (from 1 to 13) for each criterion. MeSH: Medical Subject Headings.



Discussion

Principal Findings

The overall objective of this paper is to present the value of using the user testing method to collect both quantitative and qualitative data, and to actively involve end users through an example of the evaluation of sorting formulae. A total of 2 formulae for ranking search results were evaluated to design a ranking formula that met the needs and expectations of the GPs and hospital registrars.

The first part of the study (scoring the relevance of both formulae's search results) enabled us to compare the respective ranking efficiencies and participant preferences. Of the 20 users, 14 (70%) preferred formula B. These users liked articles that formed a consensus in the field, that is, reviews and meta-analyses that; for example, contained peer-approved definitions. Case reports and publications on highly specific

elements (eg, a subcategory of patients) were judged to be of little use in practice. Of the 20 participants, 4 (20%; 3 GPs and 1 registrar) preferred formula A; the search results were more recent (ie, covering the latest theories or discoveries), although the level of evidence was lower. It is noteworthy that the more experienced GPs tended to be less interested in general articles because the latter taught them nothing new about their practice. A single ranking formula cannot meet all possible needs and expectations, so the user should be given a means to personalize the interface [37].

The second part of our study generated data on the perceived importance of the ranking criteria. Unquestionably, the ranking criteria judged by the participants to be the most valuable were those related to text data, that is, the match between the article's metadata and the user's search query. Thus, criteria such as *title*, *abstract*, *author keywords*, and *major MeSH terms* were often rated as the most important. The other criteria (ie, *the presence of an abstract*, *affiliation with a qualifier*, *and exploded indexing*) were judged to be useful, albeit mainly for differentiating between articles that already met the other criteria. The *type of publication* criterion was considered interesting because it highlighted the most reliable articles. Finally, the *year of publication* was a controversial criterion, the usefulness of which depended on the user's search purpose.

In view of the results, formula B was selected for further study. Owing to our analysis of the verbalizations given during the formulae and criteria ranking steps, we were able to produce several adaptations of this ranking formula to better meet the users' needs (Table 7). Each proposal was discussed between the Clinical Investigation Centre for Innovative Technology evaluation team and the design team of D2IM until they eventually reached a consensus. The opportunities for improvement, justifications, and state of implementation are listed in Table 7.

The main changes were the addition of the *keyword in the abstract* criterion and modification of the *type of publication* criterion. In addition to promoting consensus articles in a given field (eg, meta-analyses and literature reviews), the *type of publication* criterion downgrades articles that have little informative or scientific value for the reader (errata, questions and answers, personal narratives, etc). Another improvement discussed with the project partners was to provide the option to personalize LiSSa's relevance ranking formula; for example, by manually adding customized search criteria or by automatically learning from users' data to determine their preferences. Another approach for creating ranking formulae is machine learning based on the analysis of large quantities of user data [14]. Ultimately, it would be interesting to assemble this type of data for LiSSa and thus determine whether machine learning–based ranking formulae would differ from the formulae assessed in this study.

Table 7. Adaptations of the ranking formula, the associated justifications, and their state of implementation.

Criterion	Opportunity for improvement	Justification	State of implementation
Abstract	Take account of the keyword's presence in the abstract.	Currently, the abstract is not considered at all, even though (on average), it was the second most important criterion, right after the title. However, the abstract is less strictly controlled than the author keywords and the MeSH[a] indexing, giving it less weight that the latter.	In total, 3 points have been attributed to this criterion.
Subtitle	Lower is the weight attributed to the keyword's presence in the subtitle, relative to its presence in the title.	The subtitle had the same weighting as the title (ie, 10) but was judged to be less important by the participants because it was less useful.	The number of points attributed to this criterion has dropped from 10 to 8.
Type of publication	Add a subcriterion to downgrade types of publication judged to be irrelevant by the users.	The *type of publication* criterion in formula B favors certain types of publication. The users recommended downgrading the types of publication of little practical interest for the users (eg, editorials, errata, historical articles, and letters).	A subcriterion had been added to the *type of publication* criterion. As well as awarding 3 points to certain publications, it removed 1 point for errata, questions and answers, personal accounts, portraits, commentaries, historical articles, editorials, letters, and case reports.
Associated with a qualifier	Promote subject headings without a qualifier, except when the keyword is a qualifier.	To prioritize articles that generally address the search subject in first search results, adding the *associated with a qualifier* criterion was recommended. Thus, subject headings without a qualifier will be favored, except when the qualifier is also one of the user's keywords.	One point is added when the subject heading is not associated with a qualifier.
The journal's impact	Add this criterion but do not give it much weight.	This criterion is not of major importance to users but can be useful for differentiating between 2 articles with the same score. It was recommended that this criterion should be taken into account when calculating the scores but should not be given much weight.	This item has not yet been incorporated into the LiSSa database. At present, this information is available for only 30% of articles; it will therefore be necessary to determine the relevance of integrating this criterion into the formula.
Operation of the ranking formula	Add the points awarded for the *title* and m*ajor or minor MeSH term* criteria.	During the tests, some publications considered by the participants to be off-topic were listed in the top search results (eg, an article on bipolar depression for the query on *treatment-resistant depression*). To limit the risk of seeing off-topic publications in the top search results, it was recommended to add points awarded for the *title* and *major or minor MeSH term* criteria.	This recommendation needs to be tested because it might have a negative effect on ranking the search results; it might overprioritize the articles with a large number of indexed keywords (>20, in some cases), relative to articles with few keywords.

[a]MeSH: Medical Subject Headings.

Strengths and Limitations

The main advantage of formative assessment through user testing is that the study data are useful in the design process. The first step (where participants were asked to score the search results) enabled them to become familiar with the types of articles suggested by LiSSa and to think of the criteria that were important to them when judging an article's relevance. As the rating of an article had to be justified, users had to become conscious of their judgment criteria. This first step also enabled us to identify the strengths and weaknesses of existing formulae. In step 2, the criteria ranking and the participants' justifications helped us determine which criteria most strongly influenced the target users' perception of relevance. When coupled with the strengths and weaknesses detected in step 1, these data enabled us to adjust the ranking formula's criteria, and thus to develop a formula that should better meet health care professionals' needs. This combined methodology allowed us to evaluate the formulae's performance, collect user needs and habits, and evaluate the relevance of the articles found by the search engine.

In a user-centered design process, iterative evaluation during the design phase helps improve the tool before the final evaluation [18]. In contrast to more common methods [4,8], user testing is a relevant way to look for user inputs in the design of the article ranking formula. Typical methods for evaluating search engine relevance generally compare several ranking formulae; for example, a new ranking method against an old one [14,15]. In these 2 articles, the method used involved a large number of judges, which ensured good robustness of the results. The aim of comparing 2 ranking methods was achieved, but no additional information was obtained to improve the relevance of the results. These methods did not capture the reasons why the search results were judged to be more relevant by the participants or the criteria that participants used to assess relevance. Collecting and analyzing participants' verbalizations during user testing allowed us to understand the strengths and

weaknesses of the tested formulae and to look for improvements suggested directly by end users. Even if this method is applied here in the context of a French language health scientific literature search engine, it can be used for any type of ranking formula.

Nevertheless, this study had some limitations. Formative assessments generate data to improve a formula's design but do not validate a formula per se. Several criteria must be fulfilled for reliable and robust validation: the size of the test collection, the number of judges, the number of queries, and so on, [7,38] which a formative evaluation cannot fulfill. In total, 10 GPs and 10 registrars participated in this study. Moreover, as LiSSa is already on the web, user feedback shows that other health care professionals are using the tool (nurses, specialist physicians, physiotherapists, etc): different users might have different needs. The results are not generalizable and do not validate the formula. Therefore, a larger-scale evaluation with a larger number of participants and a broader range of user profiles is needed to evaluate and validate the final version of the formula.

Finally, a significant limitation of all approaches aimed at improving search result ranking formulae relates to the quality and availability of metadata. The level of topical relevance does not depend solely on identification and weighting; for each criterion, the data must be tagged for each article in the database. If this is not performed, the addition of a criterion may have very little impact on the result ranking or may even degrade the quality of the result ranking. During this evaluation, the criteria used were based on metadata available in the LiSSa database. We can assume that a change in the available metadata would have opened new opportunities for ranking search results, and therefore, would have impacted the study results. During our tests, we asked the participants whether criteria other than those presented should be added. Several criteria had been suggested (ie, *type of population*, *medical specialty*, *methodology*, etc), but none can yet be considered for inclusion because the related metadata are not available or are not of high enough quality or coverage for reliable incorporation in the formula. This was also the case for the criteria presented during the study. Some criteria presented during the second step of the study were interesting to users but could not be implemented directly after the evaluation. For example, *journal importance* metadata were available for >30% of the articles included in the LiSSa database, which prevented us from implementing this criterion. Nevertheless, the identification of new criteria has challenged the design team to add metadata and complete the formula using new criteria in the near future.

The LiSSa database contains various types of publications from thousands of journals and hundreds of publications. Each publisher has its own rules for tagging articles because they do not address the same indexing objectives [39]. Therefore, the creation of a database of over a million articles is already a challenge, particularly with regard to harmonizing metadata of different types and formats. Thus, the reuse of metadata is an objective, within which the ranking of results is just one of many challenges. This reveals the need for a true debate among all stakeholders (ie, publishers, institutions, and users) about standardized indexing that meets various objectives (eg, ranking and archiving).

Conclusions

To conclude, LiSSa is a tool intended for practitioners who are not specialized in scientific research and who do not speak English. This study highlights the need of these end users to improve the topical relevance of the first top-ranked results. The assessment of the LiSSa search engine's result ranking formulae enabled us to draw a list of recommendations for a ranking formula that would meet the ranking needs of GPs and hospital registrars. In the next step of the project, we will assess the relevance and appropriateness of the redesigned ranking formula with regard to user needs and expectations. To this end, we shall conduct tests with a new panel of users that includes more types of health care professionals.

Acknowledgments

The authors thank all general practitioners and registrars who voluntarily participated in this evaluation.

The LiSSa project was partly funded by the ANR TecSan program as part of the *Base de Données Bibliographiques en français* (French Bibliographic Database) project (reference: BDBfr, ANR-14-CE17-0020).

The maintenance of the database and website is supported by funding from the Health Professionals' Press and Publishing Union (*Syndicat de la Presse et de l'Edition des Professionnels de Santé*).

Conflicts of Interest

None declared.

References

1. Delestre N, Malandain N. Du Web des documents au Web sémantique [From the Document Web to the Semantic Web]. Bois-Guillaume: KLOG; 2017.
2. Landhuis E. Scientific literature: information overload. Nature 2016 Jul 21;535(7612):457-458. [doi: 10.1038/nj7612-457a] [Medline: 27453968]
3. Nabeel Asim M, Wasim M, Usman Ghani Khan M, Mahmood W. Improved biomedical term selection in pseudo relevance feedback. Database (Oxford) 2018 Jan 01;2018:bay056 [FREE Full text] [doi: 10.1093/database/bay056] [Medline: 29982558]

4. Elbedweihy K, Wrigley S, Clough P, Ciravegna F. An overview of semantic search evaluation initiatives. J Web Semantics 2015 Jan;30:82-105 [FREE Full text] [doi: 10.1016/j.websem.2014.10.001]
5. Robertson S. On the history of evaluation in IR. J Inf Sci 2008 Jun 13;34(4):439-456 [FREE Full text] [doi: 10.1177/0165551507086989]
6. Vicedo J, Gómez J. TREC: experiment and evaluation in information retrieval. J Am Soc Inf Sci 2007 Apr;58(6):910-911 [FREE Full text] [doi: 10.1002/asi.20583]
7. The philosophy of information retrieval evaluation. In: Evaluation of Cross-Language Information Retrieval Systems. Berlin, Heidelberg: Springer; 2002.
8. Kelly D, Sugimoto C. A systematic review of interactive information retrieval evaluation studies, 1967-2006. J Am Soc Inf Sci Tec 2013 Mar 01;64(4):745-770 [FREE Full text] [doi: 10.1002/asi.22799]
9. Saracevic T. Relevance: a review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance. J Am Soc Inf Sci 2007 Nov;58(13):1915-1933 [FREE Full text] [doi: 10.1002/asi.20682]
10. Borlund P, Ingwersen P. Measures of relative relevance and ranked half-life: performance indicators for interactive IR. In: Proceedings of the 21st Annual ACM/SIGIR International Conference on Research and Development in Information Retrieval. 1998 Presented at: SIGIR98: 21st Annual ACM/SIGIR International Conference on Research and Development in Information Retrieval; Aug 24 - 28, 1998; Melbourne Australia URL: https://doi.org/10.1145/290941.291019 [doi: 10.1145/290941.291019]
11. Harter SP. Psychological relevance and information science. J Am Soc Inf Sci 1992 Oct;43(9):602-615 [FREE Full text] [doi: 10.1002/(sici)1097-4571(199210)43:9<602::aid-asi3>3.0.co;2-q]
12. Turpin AH, Hersh W. Why batch and user evaluations do not give the same results. In: Proceedings of the 24th ACM/SIGIR International Conference on Research and Development in Information Retrieval. 2001 Presented at: SIGIR01: 24th ACM/SIGIR International Conference on Research and Development in Information Retrieval; 2001; New Orleans Louisiana USA URL: https://doi.org/10.1145/383952.383992 [doi: 10.1145/383952.383992]
13. Sanderson M, Paramita ML, Clough P, Kanoulas E. Do user preferences and evaluation measures line up? In: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. 2010 Presented at: SIGIR '10: The 33rd International ACM SIGIR conference on research and development in Information Retrieval; Jul 19 - 23, 2010; Geneva Switzerland URL: https://doi.org/10.1145/1835449.1835542 [doi: 10.1145/1835449.1835542]
14. Fiorini N, Canese K, Starchenko G, Kireev E, Kim W, Miller V, et al. Best Match: new relevance search for PubMed. PLoS Biol 2018 Aug;16(8):e2005343 [FREE Full text] [doi: 10.1371/journal.pbio.2005343] [Medline: 30153250]
15. Ji Y, Ying H, Tran J, Dews P, Massanari RM. Integrating unified medical language system and association mining techniques into relevance feedback for biomedical literature search. BMC Bioinformatics 2016 Jul 19;17 Suppl 9:264 [FREE Full text] [doi: 10.1186/s12859-016-1129-z] [Medline: 27453982]
16. Bruza P, Chang V. Perceptions of document relevance. Front Psychol 2014;5:612 [FREE Full text] [doi: 10.3389/fpsyg.2014.00612] [Medline: 25071622]
17. Lee SS, Theng YL, Goh DH, Foo SB. Subjective relevance: implications on digital libraries for experts and novices. In: Digital Libraries: International Collaboration and Cross-Fertilization. Berlin, Heidelberg: Springer; 2004.
18. Ergonomics of human-system interaction — Part 210: human-centred design for interactive systems. ISO. URL: https://www.iso.org/standard/77520.html [accessed 2022-02-27]
19. Kolski C. Analyse et conception de l'IHM: interaction homme- machine pour les SI. Volume I. Paris, France: Hermes Science Publications; 2001.
20. Carayon P, Hoonakker P. Human factors and usability for health information technology: old and new challenges. Yearb Med Inform 2019 Aug;28(1):71-77 [FREE Full text] [doi: 10.1055/s-0039-1677907] [Medline: 31419818]
21. Kushniruk A, Nohr C, Borycki E. Human factors for more usable and safer health information technology: where are we now and where do we go from here? Yearb Med Inform 2016 Nov 10(1):120-125 [FREE Full text] [doi: 10.15265/IY-2016-024] [Medline: 27830239]
22. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. Clin Psychol (New York) 2016 Jun;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
23. Wiklund ME, Kendler J, Strochlic AY. Usability Testing of Medical Devices. Boca Raton, FL: CRC Press; 2016.
24. Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
25. Littérature Scientifique en Santé homepage. Littérature Scientifique en Santé. URL: https://www.lissa.fr/dc/#env=lissa [accessed 2022-02-27]
26. Griffon N, Rollin L, Schuers M, Douze L, Delerue D, Dutoit D, et al. LiSSa, Littérature Scientifique en Santé : une base de données bibliographique en français [LiSSa, health scientific literature: a French bibliographic database]. Rev Prat 2017 Feb;67(2):134-138. [Medline: 30512844]
27. Darmoni SJ, Leroy JP, Baudic F, Douyère M, Piot J, Thirion B. CISMeF: a structured health resource guide. Methods Inf Med 2000 Mar;39(1):30-35. [Medline: 10786067]

28. Grosjean J, Merabti T, Dahamna B, Kergourlay I, Thirion B, Soualmia LF, et al. Health multi-terminology portal: a semantic added-value for patient safety. Stud Health Technol Inform 2011;166:129-138. [Medline: 21685618]
29. Griffon N, Schuers M, Soualmia LF, Grosjean J, Kerdelhué G, Kergourlay I, et al. A search engine to access PubMed monolingual subsets: proof of concept and evaluation in French. J Med Internet Res 2014 Dec 01;16(12):e271 [FREE Full text] [doi: 10.2196/jmir.3836] [Medline: 25448528]
30. Goswami P, Gaussier E, Amini M. Exploring the space of information retrieval term scoring functions. Inf Process Manag 2017 Mar;53(2):454-472 [FREE Full text] [doi: 10.1016/j.ipm.2016.11.003]
31. Jiang JY, Liu J, Lin CY, Cheng PJ. Improving ranking consistency for web search by leveraging a knowledge base and search logs. In: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. 2015 Presented at: CIKM'15: 24th ACM International Conference on Information and Knowledge Management; Oct 18 - 23, 2015; Melbourne Australia URL: https://doi.org/10.1145/2806416.2806479 [doi: 10.1145/2806416.2806479]
32. Agichtein E, Brill E, Dumais S. Improving web search ranking by incorporating user behavior information. In: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. 2006 Presented at: SIGIR06: The 29th Annual International SIGIR Conference; Aug 6 - 11, 2006; Seattle Washington USA URL: https://doi.org/10.1145/1148170.1148177 [doi: 10.1145/3308774.3308778]
33. Darmoni SJ, Pereira S, Sakji S, Merabti T, Prieur E, Joubert M, et al. Multiple terminologies in a health portal: automatic indexing and information retrieval. In: Artificial Intelligence in Medicine. Berlin, Heidelberg: Springer; 2009.
34. Järvelin K, Kekäläinen J. Cumulated gain-based evaluation of IR techniques. ACM Trans Inf Syst 2002 Oct;20(4):422-446 [FREE Full text] [doi: 10.1145/582415.582418]
35. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
36. Krippendorff K. Content Analysis An Introduction to Its Methodology. Thousand Oaks, California: SAGE Publications Inc; 2004.
37. Teevan J, Dumais S, Horvitz E. Beyond the commons: investigating the value of personalizing web search. In: Proceedings PIA 2005: Workshop on New Technologies for Personalized Information Access. 2005 Presented at: PIA 2005: Workshop on New Technologies for Personalized Information Access; Jul 24-30, 2005; Edinburgh, UK.
38. Yang Z. Improving the reliability and robustness of information retrieval evaluation. University of Melbourne. 2019. URL: https://minerva-access.unimelb.edu.au/handle/11343/233322 [accessed 2022-02-27]
39. Gregg W, Erdmann C, Paglione L, Schneider J, Dean C. A literature review of scholarly communications metadata. Res Ideas Ootcome 2019 Aug 05;5:e38698 [FREE Full text] [doi: 10.3897/rio.5.e38698]

Abbreviations

D2IM: Department of Medical Informatics and Information
GP: general practitioner
MeSH: Medical Subject Headings
NDCG: normalized discounted cumulative gain
NLM: National Library of Medicine

Edited by A Kushniruk; submitted 11.05.21; peer-reviewed by J Griffith, R Ologeanu-Taddei; comments to author 28.06.21; revised version received 07.07.21; accepted 09.01.22; published 25.03.22.

Please cite as:
Douze L, Pelayo S, Messaadi N, Grosjean J, Kerdelhué G, Marcilly R
Designing Formulae for Ranking Search Results: Mixed Methods Evaluation Study
JMIR Hum Factors 2022;9(1):e30258
URL: https://humanfactors.jmir.org/2022/1/e30258
doi:10.2196/30258
PMID:35333180

Original Paper

Comparing International Experiences With Electronic Health Records Among Emergency Medicine Physicians in the United States and Norway: Semistructured Interview Study

Gracie Garcia[1], BA; Christopher Crenner[1], MD, PhD

Department of History and Philosophy of Medicine, University of Kansas School of Medicine, Kansas City, KS, United States

Corresponding Author:
Gracie Garcia, BA
Department of History and Philosophy of Medicine
University of Kansas School of Medicine
3901 Rainbow Blvd
Kansas City, KS, 66160
United States
Phone: 1 6209514171
Email: grace.c.garcia1@gmail.com

Abstract

Background: The variability in physicians' attitudes regarding electronic health records (EHRs) is widely recognized. Both human and technological factors contribute to user satisfaction. This exploratory study considers these variables by comparing emergency medicine physician experiences with EHRs in the United States and Norway.

Objective: This study is unique as it aims to compare individual experiences with EHRs. It creates an opportunity to expand perspective, challenge the unknown, and explore how this technology affects clinicians globally. Research often highlights the challenge that health information technology has created for users: Are the negative consequences of this technology shared among countries? Does it affect medical practice? What determines user satisfaction? Can this be measured internationally? Do specific factors account for similarities or differences? This study begins by investigating these questions by comparing cohort experiences. Fundamental differences between nations will also be addressed.

Methods: We used semistructured, participant-driven, in-depth interviews (N=12) for data collection in conjunction with ethnographic observations. The conversations were recorded and transcribed. Texts were then analyzed using NVivo software (QSR International) to develop codes for direct comparison among countries. Comprehensive understanding of the data required triangulation, specifically using thematic and interpretive phenomenological analysis. Narrative analysis ensured appropriate context of the NVivo (QSR International) query results.

Results: Each interview resulted in mixed discussions regarding the benefits and disadvantages of EHRs. All the physicians recognized health care's dependence on this technology. In Norway, physicians perceived more benefits compared with those based in the United States. Americans reported fewer benefits and disproportionally high disadvantages. Both cohorts believed that EHRs have increased user workload. However, this was mentioned 2.6 times more frequently by Americans (United States [n=40] vs Norway [n=15]). Financial influences regarding health information technology use were of great concern for American physicians but rarely mentioned among Norwegian physicians (United States [n=37] vs Norway [n=6]). Technology dysfunctions were the most common complaint from Norwegian physicians. Participants from each country noted increased frustration among older colleagues.

Conclusions: Despite differences spanning geographical, organizational, and cultural boundaries, much is to be learned by comparing individual experiences. Both cohorts experienced EHR-related frustrations, although etiology differed. The overall number of complaints was significantly higher among American physicians. This study augments the idea that policy, regulation, and administration have compelling influence on user experience. Global EHR optimization requires additional investigation, and these results help to establish a foundation for future research.

(JMIR Hum Factors 2022;9(1):e28762) doi:10.2196/28762

KEYWORDS

electronic health records; electronic medical records; health information technology; health information exchange; health policy; international; emergency medicine; medical informatics; meaningful use; burnout

Introduction

Background

Correlations between electronic health records (EHRs) and physician frustrations have been well described throughout informatics literature. The phenomenon of high user dissatisfaction is often attributed to increased administrative requirements, decreased face-to-face patient time, information overload, and limited interoperability [1-12]. This technology has been analyzed on both local and global scales [13-17]; however, few studies have compared users from different countries who practice in parallel clinical settings. Our study compares emergency medicine (EM) physician experiences with EHRs in the United States and Norway. In addition, societal and cultural differences are carefully considered while analyzing components that may affect user satisfaction.

Health information technology (HIT) is used in many countries, but deployment of EHRs vary [18-23]. Despite global HIT use, the United States is perhaps the most prominent generator of informatics research that emphasizes the shortcomings of this technology. Currently, there are few studies that consider or compare international EHR experiences. However, a recent study by Downing et al [24] found that even when using the same vendor (Epic Systems), American physicians had significantly longer documentation and were less likely to report satisfaction or improved work efficiency compared with those in Australia and Singapore. Our results contribute to this small body of research.

Objectives and Measured Outcomes

This study considers factors that contribute to EHR user satisfaction by comparing individuals with similar professional responsibilities in different national contexts. Are the negative consequences of this technology shared among countries? Does it affect medical practice? What determines user satisfaction? Can this be measured between international cohorts? Are there specific factors that account for similarities or differences? This study explores these questions while considering influential variables from a sociopolitical–technological context.

Primary outcomes include the overall EHR experience and specific opinions within each cohort. This was achieved by conducting structured interviews in the hospital and observing behaviors within the physician's typical working environment. Thematic analysis allowed the quantification and comparison of common topics. Notable differences may help to identify targeted solutions for HIT optimization. For example, if users from each cohort believe that software interfaces are challenging to use, it could indicate that technology-specific factors (understanding and using computers) significantly increase frustration. On the other hand, differences may identify solutions that may have otherwise been overlooked.

Secondary outcomes assess participant responses within a sociocultural context, as HIT infrastructure differs among countries [16-23,25]. Previous research shows that successful EHR use is greatly influenced by social and governmental constructs [9,24,26,27]. A general understanding of the current HIT status and health care infrastructure in the United States and Norway supports the interpretation of the data. We briefly discuss this information before proceeding.

United States: Emergency Care and Current EHR Status

Since the Emergency Medical Treatment and Labor Act of 1986, hospitals must provide consultation, screening examination, ancillary testing, and stabilization of anyone concerned with a life-threatening condition, regardless of their ability to pay [28-30]. Patients are evaluated in the emergency department (ED) after arriving via ambulance, private vehicle, or walking in. Physicians who staff the ED receive formal EM training by completing an EM residency for 3-4 years following medical school [31]. Although the ED functions as a hospital's gatekeeper, studies show that only a small number of ED visits result in admission [32]. National increase in low-acuity ED patient volume has been attributed to multiple factors including rising health care costs, primary care shortages, and lack of access to after-hour care [33-37].

Integration of technology and health care started in 2004 with the establishment of the Office of the National Coordinator for HIT, but widespread EHR use did not occur until after the HIT for Economic and Clinical Health Act was passed in 2009 [38]. This legislation provided monetary incentives for government-certified EHR adoption and implementation [39]. Pressure for rapid health care digitization generated numerous unintended consequences including industrial arms race that many policy makers did not consider [40-43]. As of 2017, 96% of hospitals in the United States had implemented technology certified by the Department of Health and Human Services [44].

The 21st Century Cures Act prohibits companies and organizations from intentionally restricting health information exchange (HIE) capabilities for monetary benefit [45]. Nevertheless, evidence indicates that *information blocking* still occurs in the United States [46,47], and a 2018 Report to Congress showed only 51% of hospital physicians had electronic access to necessary patient information from other facilities at the point of care [44]. The private sector has been meeting interoperability demands as evidenced by programs like Epic's Care Everywhere [27]. However, the extent of clinical data availability is dependent on the participating facilities [48]. In 2018, the US Department of Veterans Affairs (VA) announced partnership with Cerner, an EHR company that will eventually be the sole vendor to all VA facilities that serve military populations [49].

Norway: Emergency Care and Current EHR Status

Inpatient and specialist care are provided by state-owned hospitals and managed by 4 geographically distinct government subdivisions known as Regional Health Authorities [50,51]. A

total of 428 local municipalities are responsible for supplying primary care including after-hour access [50]. Municipalities have urgent care centers with on-call physicians (*legevakt*) [50]. The ED or *acute receiving area* (*akuttmottak*) is only accessible via ambulance or physician referrals [50,52]. The department is traditionally staffed by internal medicine, neurology, orthopedics, and surgery physicians [53]; however, EM was recently recognized as an independent specialty in Norway in 2017 [52]. Historically, ambulance and other health personnel would communicate with hospitals to determine the most appropriate inpatient specialty service to receive the patient upon arrival [50].

Medical records from hospitals and outpatient facilities are not integrated, but messaging systems embedded within EHR software allow providers to collaborate [51]. In 2008, the government recognized the interoperability needs and launched a national HIE platform in 2012 known as *Core Journal* (*Kjernejournalen*) [54]. This gives all Norwegian physicians access to critical patient information, regardless of where previous treatment was provided [51,54]. It includes data necessary to prevent unfavorable outcomes that may be difficult to obtain during emergency situations such as severe allergies, ongoing treatments (eg, dialysis), rare serious conditions (eg, hemophilia), and medications dispensed at any Norwegian pharmacy [55]. Research shows that the most used function is the pharmaceutical tracking tool as it provides up-to-date medication information without additional manual data-entry requirements from physicians [56,57].

In 2013, the Directorate of Health recommended the integration of all eHealth and developed the initiative *One Patient–One Record* (*Én innbygger–Én journal*) [58,59]. In 2019, a US $296 million contract was signed with an American EHR company (Epic Systems) to eventually function as the nation's sole HIT supplier [13,60]. The pilot program *Health Platform* (*Helseplattformen*) is scheduled to launch during the spring of 2022 in Central Norway, 1 of the 4 Regional Health Authorities [61]. Current studies indicate optimistic expectations mixed with concern as protected health information will eventually be exchanged across administrative, geographical, and institutional boundaries [62]. Regional governments created *consensus groups* comprised of health care professionals from >80 municipalities that are involved in software configuration and design [63]. After implementation, community physicians and analysts will continue to optimize the functionality for regional and practice needs, whereas Epic Systems will be involved to a lesser extent [13].

Methods

Participants and Setting

This study was conducted at the University of Kansas Medical Center (KUMC) in Kansas City, Kansas, and at the Akershus University Hospital in the Lørenskog municipality outside of Oslo. Bed capacity at each hospital was approximately 1000 beds [58,59]. Recruitment emails were sent to physicians involved in acute care at these facilities. In the United States, participants were board-certified EM physicians, whereas in Norway, participants were surgeons who provide services within the *akuttmottak*. A total of 12 interviews were conducted, 6 (50%) at each location. Average conversation lengths were 39.1 (SD 15.8) minutes.

Data Collection

Data collection included face-to-face semistructured interviews and environmental observations. This was possible by conducting each interview on site at the hospitals. Participants were willing to show the typical documentation and clinical workflow to the interviewer (GG). This was essential when collecting Norwegian data, as the interviewer had no previous first-hand experience with this health care system. This provided context when participants referred to specifics of the EHR. Without this background the contextual understanding of participants' answers would have been severely limited. All the interviews were conducted in English, as all the participants were proficient in this language. Conversations were audio-recorded on a passcode-protected device and then transcribed for further analysis. Privacy was retained by deidentifying the participants.

After obtaining written informed consent, standardized questions were used to obtain the following information from each participant: (1) demographics, (2) cultural and individual values, (3) individual comfort with general technology, (4) previous record experiences (electronic or paper), (5) observations of colleagues regarding EHR use, (6) individual attitudes toward EHR at current facility, (7) perceived usability (intuitive interfaces, software functionality, interoperability, workflow efficiencies, and centralized data repository), and (8) how the technology has shaped individual practice.

Follow-up questions varied based on individual responses. Participants were also asked about their knowledge, opinions, or questions regarding the other cohort's electronic health care infrastructure. Natural conversation flow permitted additional discussion, allowing deeper exploration of ideas as they appeared organically. Additional questions developed throughout data collection were based on previous participant answers and cumulative observations. For example, US interviews were completed first and the responses involved specific negative consequences of the EHR without prompting. If these topics were never mentioned by the Norwegians, the interviewer inquired about them directly at the end of the discussion.

To conclude each interview, participants were asked if they had specific questions for the physicians in the other country. Following data collection, questions and answers were distributed to participants in addition to the background information on each country's health care system. This allowed deeper understanding of individual perceptions while generating rich discussion. In addition, participants in Norway were explicitly asked about *Kjernejouralen* use. This study was reviewed by and received institutional review board approval from KUMC while abiding by the General Data Protection Regulation.

Analysis

The US interviews were completed first, followed by interviews in Norway. Using grounded theory, themes emerged and evolved

throughout the entire data collection process. As no single method captured the complexities of these data, analysis triangulation was necessary. First, transcripts underwent numerous thematic analyses to identify patterns between the cohorts. This was the initial formal approach to derive meaning from the vast and rich collected data. Similar to grounded theory, this exploratory methodology allows continuous hypothesis development throughout analysis progression. Narrative analysis was conducted to provide further insight into the mindset, perspectives, and attitudes toward EHRs. In addition, direct quotes were used to support the findings and may help the reader appreciate the nuances of the social context and emotion.

Early in the analysis process, 2 broad themes were identified—*perceived EHR benefits* and *perceived EHR disadvantages*. To gain deeper understanding of the data, interpretive phenomenological analysis and simple content analysis were used. Both methods aid in succinctly summarizing concepts based on individual experiences while providing some quantitative comparison. These techniques paired with the NVivo software (QSR International) helped to distinguish conceptual patterns between the cohorts, and ultimately resulted in the construction of the following 4 main code groups: *US perceived EHR benefits, Norway perceived EHR benefits, US perceived EHR disadvantages, and Norway perceived EHR disadvantages*.

Transcriptions were analyzed using the NVivo software (QSR International). As the perceived EHR benefits or disadvantages were found within the text, they were assigned to 1 of the 4 code groups based on context and cohort. The NVivo word frequency and query search functions were used to generate categories within the encoded text to enrich the results. The software allowed searches to include exact word matches, stemmed words, and synonyms. The search criteria details are presented in Textboxes 1 and 2. Identical queries regarding perceived EHR benefits and disadvantages were conducted for both cohorts. Query results were analyzed and refined to ensure that the terms were not taken out of context. The total number of results for each group was tabulated and compared. This is displayed in Figure 1. Comparing the categorical patterns provides concrete examples of varying priorities, opinions, and perspectives from the 2 cohorts. In addition, it examines the advantages and flaws of HIT implementation within each health care system.

Textbox 1. Categories of perceived electronic health record benefits.

Category and search criteria (include exact matches, stemmed words, and synonyms)

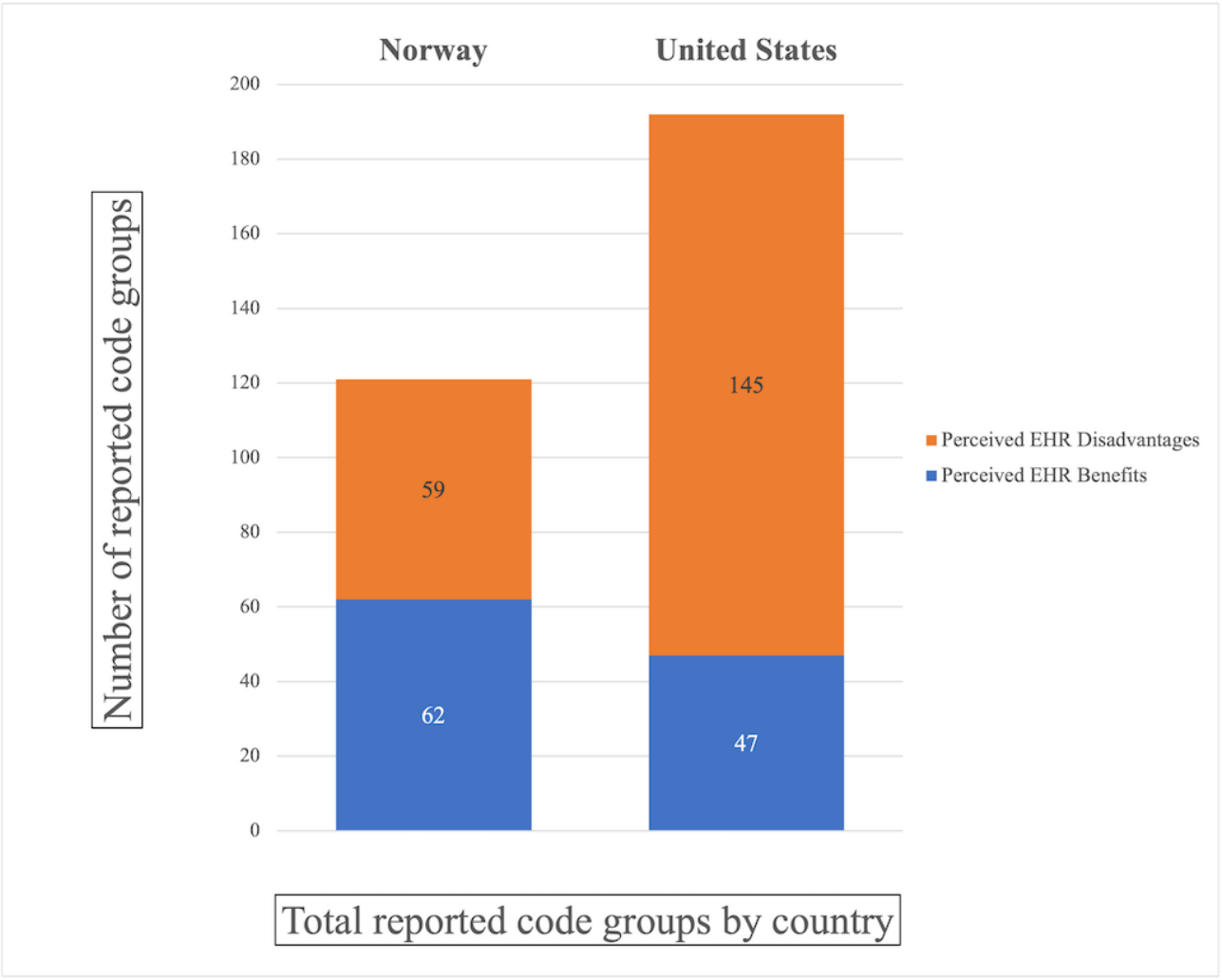
- Patient safety and improved care: *safety*, *benefit*, *care*, *improve*, *alert*, *allergy*, *interaction*, *medication*, *automated*, *error*, *writing*, *legible*, and *mistakes*
- Access to useful clinical information: *accessibility*, *information*, *view*, *records*, *journal*, *chart*, *review*, *report*, *previous*, *tracking*, *results*, *history*, *important*, *critical*, *clinical*, *diagnosis*, *remote*, *exchange*, *facility*, *interoperable*, *capability*, *cloud*, *electronic*, *time*, and *speed*
- Data organization: organization, *sort*, *filter*, *search*, *usability*, *function*, *central*, *record*, *history*, *chart*, *journal*, *ease*, *efficient*, and *review*
- Enhanced communication: *communication*, *interaction*, *order*, *results*, *review*, *chart*, *record*, *journal*, *information*, *patient*, *encounter*, and *clarification*

Textbox 2. Categories of perceived electronic health record disadvantages.

Category and search criteria (include exact matches, stemmed words, and synonyms)

- Excessive or irrelevant data: *excess*, *irrelevant*, *overload*, *quantity*, *redundant*, *limit*, *volume*, *amount*, *organize*, *filter*, *lost*, *search*, *data*, *clinical*, *benefit*, and *documentation*
- Poor interoperability: *interoperability*, *access*, *view*, *restrict*, *facility*, *exchange*, *data*, *information*, *chart*, *journal*, *record*, *outside*, *cloud*, *capability*, *hospital*, and *clinic*
- Increased workload: *work*, *workload*, *time*, *hour*, *administrative*, *requirement*, *documentation*, *efficient*, *amount*, *burden*, *click*, *task*, and *clerical*
- Software complexities: *software*, *complex*, *interface*, *intuitive*, *user*, *difficult*, *friendly*, *usability*, *navigate*, *understand*, *function*, *options*, *programs*, *system*, *load*, *slow*, *lag*, *ease*, *options*, *orders*, *run*, and *technology*
- Hardware malfunctions: *hardware*, *malfunction*, *crash*, *process*, *failure*, *update*, *IT*, *program*, *develop*, *technology*, *support*, *computer*, *device*, *speed*, and *paper*
- Financial influence: *financial*, *money*, *reimbursement*, *billing*, *profit*, *cost*, *incentive*, *code*, *dollars*, *relative value unit or RVU*, *regulation*, *mandate*, *clinical*, *value*, *price*, *payment*, and *business*

Figure 1. Total number of reported code groups by country. EHR: electronic health record.



This process was conducted by the interviewer for retained consistency while considering abstract factors including nonverbal communication, clinical environment, and cultural norms. This technique was repeatedly used to explore topic relationships, consider causality, and help find thematic saturation within the populations.

Results

Overview

All the participants described both pros and cons of their experience with the EHRs and both groups agreed that modern medicine is heavily dependent on this technology. In general, Norwegian physicians had a slight propensity to report benefits (62 total perceived benefits reported) compared with disadvantages (59 total perceived disadvantages reported). In contrast, the American cohort frequently expressed unfavorable perceptions, reporting 145 total perceived disadvantages and only 47 total perceived benefits. These results are summarized in Figure 1.

Perceived Benefits

Access to relevant patient information was the most commonly reported benefit in both countries. This included viewing previous diagnostic studies, clinical notes, and laboratory results. The Norwegian physicians were 1.7 times more likely to refer to these specific benefits (Norway [n=33] vs United States [n=20]). When American physicians mentioned this tool, they often also noted significant limitations owing to poor interoperability between competing HIT supply companies and health care facilities. A commonly perceived positive EHR outcome in both cohorts was improved patient safety. The results were moderately comparable between the 2 countries with American physicians referencing patient safety 20 times and Norwegian physicians referencing patient safety 15 times. An example that was frequently mentioned by participants was the automated alerts about patient allergies or drug–drug interactions. Many also believed that it has decreased unnecessary errors caused by illegible handwriting.

Perceived Disadvantages

In general, there was a much broader range of topics related to perceived disadvantages when compared with benefits. The belief that EHRs have increased physicians' workload was common to both cohorts. However, this was mentioned 2.6 times more frequently by the Americans (United States [n=40] vs Norway [n=15]). The most reported disadvantage was how increased clerical work detracted from efficiency. American physicians also discussed that they believe the required documentation has minimal, if any, clinical utility.

In Norway, the 2 most frequently discussed disadvantages of EHRs included software complexities (Norway [n=36] vs United States [n=25]) and hardware malfunctions (Norway [n=15] vs United States [n=7]). Every other disadvantage category was more common among the US cohort. In addition to increased workload, other categories included excessive and irrelevant data (United States [n=25] vs Norway [n=12]) and poor interoperability (United States [n=34] vs Norway [n=14]). The most significant difference between the cohorts was regarding the financial influence of the EHRs (United States [n=37] vs Norway [n=6]). Each American physician expressed without prompting that the primary purpose of EHRs within the United States is for billing rather than to improve patient care. This was often attributed to competing business models among HIT suppliers, insurance companies, and hospital administrations.

One disadvantage exclusive to the American cohort involved the legal implications of the EHRs. The interviewer never initiated this topic, yet it was brought up by half of the American participants. They strongly believed that the normalization of defensive medicine is a result of the society's legal climate. Despite possessing adequate medical training and clinical judgment, clinicians often feel compelled to order extensive workups to protect themselves from future prosecution. In addition, these physicians mentioned that redundant testing is routinely performed because of limited HIE among surrounding health care facilities.

Additional Observations

An interesting observation shared by both the cohorts was that their older colleagues expressed higher levels of EHR-related frustration. This was mentioned by 9 among all 12 participants—6 Norwegians and 3 Americans. These 9 individuals self-reported that they felt proficient in using technology but did not believe it influenced their own opinion of EHRs.

In addition, both Norwegians and Americans believed that the rapid processing speeds of personal devices may contribute to unrealistic EHR performance expectations. Many realized that top information technology developers are recruited to sectors outside of health care; however, they believed that usability would improve if companies such as Apple or Google developed the software:

> *It's very hard to keep up with ever-changing new technology. As you get older you don't have the stamina. Programs may also seem frustrating because they don't run as quickly as most of our personal devices.* [Norway, participant 4]

At the start of each interview, the physicians were asked how cultural values influenced personal beliefs or medical practice, as other studies have described health care systems as a reflection of national ideals [26]. This question was intended to highlight nuanced variables that exist when comparing dissimilar populations. Unsurprisingly, the participant responses revealed differing values between the countries. Responses were not superior or inferior, just different. When describing how cultural values influence their current practice of medicine, the American participants used words such as *help*, *kind*, and *caring*. Common Norwegian terms included *open-minded*, *equality*, and *empathy*. Although these results have limited application in determining EHR satisfaction, it reinforces the importance of cultural context when developing solutions for specific populations.

Discussion

Adding quantitative values to our qualitative analysis creates an overt visualization of the differences between EHR users in both countries. We have provided a more comprehensive exploration of the influencing factors.

Clerical Burden and Reimbursement

Increased administrative tasks that yield minimal patient benefit created frustration for all physicians; however, it was significantly higher among US participants. In Norway, physicians must include appropriate diagnosis or procedure codes for hospital reimbursement using the Diagnosis Related Groups system, which includes approximately 980 codes [64]. These codes generate approximately 50% of the hospital revenue, with the remaining financed from fixed government payments [65]. Norwegians are skeptical of potential changes following the national implementation of Epic Systems. The participants voiced concern regarding slowly evolving into an American health care model. Some Norwegian participants first noticed this shift after hospital reimbursement became partially integrated with diagnosis codes:

> *With new public management reform within the last 30 years, we have also noticed health care has changed to suit their needs. Most Norwegian physicians are attentive and oppose this. We also have the union (Norwegian Medical Association) who oppose it. It isn't in our immediate power to change those things and they must come from a higher level.* [Norway, participant 6]

American reimbursement is complex owing to a multi-payer system that includes government agencies, insurance companies, health maintenance organizations, employers, and individual patients [66]. Although many countries use the International Classification of Disease, the United States is one of the few countries that use it for both diagnosis and billing, while including more than 90,000 codes [67]. Compared with Norway, the United States uses the entire medical record for reimbursement. The billing *level* is determined by the quantity of the documented elements within each note section (ie, history of present illness, review of systems, and physical exam) with more elements correlating with higher billing levels, resulting in increased reimbursement [68]. Physicians must also filter through long, redundant, and confusing lists of diagnoses to choose the most detailed option [24]. Another form of reimbursement, relative value units, is also extracted from the EHR. These are based on >8000 procedure codes extrapolated to measure physician productivity, which are then used to determine department or individual reimbursement [66,69]; 1 physician shared that for the last 2 years, 25% of their salaries depended on the individual relative value units generated:

> *Institutions now look at emergency departments as revenue generators. We cost society more and in the end the patient loses directly and indirectly.* [United States, participant 3]

American documentation tends to be 4 times longer than that of other countries, without offering any additional clinical information [48]. The position of *medical scribes* (nonclinical personnel who are trained to provide documentation assistance and workflow support) was created as a possible solution to this problem. Research demonstrates that scribes are valued team members and improve provider satisfaction [70]. This sentiment was echoed by American physicians, whereas Norwegians were unfamiliar with this occupation. Gardner et al [1] showed great variety of scribe use among American specialties, with the highest use among EM physicians. This study also found that working with scribes reduced the odds of burnout by approximately 40%. They hypothesized that it was not higher because scribes are not qualified to complete certain time-consuming but physician-specific electronic tasks (eg, medication orders and in-basket management) [1]:

> *We are so opposed to these tasks that steal time we could otherwise use for clinical work. Don't you think that having a scribe is just a waste of resources? Do individual physicians actually generate enough data on a single patient that they need a scribe to help complete the documentation?* [Norway, participant 6]

Burnout

Multiple studies within the United States indicate that HIT creates undue physician burden and there is considerable correlation between high EHR frustration and burnout [1,4,7,24,43,71-76]. The 2018 National Physician Poll produced powerful data regarding how this technology affects American physicians. Only 8% of participants believed that the primary purpose of documentation is clinical, whereas 71% agreed that it significantly contributes to burnout [77]. Our study supports this argument, as many American physicians cited EHRs as a significant cause of burnout. However, these individuals clarified that it is only a single contributor to a complex and multifactorial issue:

> *I don't necessarily think the electronic aspect of EHRs are what makes them so frustrating, but rather the need of documenting in excess. When you have to do such complicated things against your will and without patient benefit, it adds to burnout rates.* [United States, participant 6]

In Norway, burnout was never mentioned spontaneously and eventually the interviewer was required to ask about it explicitly. Norwegians attributed burnout to perceived job demands, societal expectations, and degree of colleague support, which is consistent with other Norwegian studies [78-80]. EHRs were never mentioned as a source of burnout. Since 1993, Norway has conducted extensive research aimed at improving physicians' health, working conditions, and quality of life [81]. Despite burnout being less prevalent, Norway has established proactive prevention initiatives. An example is a self-referral physician counseling program and treatment facility (Villa Sana) designed to enhance coping skills and reduce emotional exhaustion [82,83].

West et al [76] considered factors that contribute to physician burnout from a global perspective. In doing so, they highlighted a previous Norwegian study that found no significant difference in burnout between physicians and other professions [84]. However, in the same study, there was a significantly increased prevalence of burnout in the United States even after adjusting for work hours and other factors [76]. Another recent US study identified systemic issues contributing to EM physician burnout. Factors include EHR limitations, long work hours, substantial educational debt, intense clinical practice, high risk of litigation, circadian rhythm disruption, chronic fatigue, blame, and isolation as a result of poor outcomes, all within the confines of an environment with zero tolerance for mistakes [85]. Our study offers informal evidence that EHRs increase burnout risk in the United States but appear noncontributory in Norway.

Core Journal (Kjernejournalen)

Of the 6 Norwegian physicians, 5 used the *Kjernejournalen* at least multiple times per week. Most information required initial manual entry, which has created additional tasks for providers. Some participants also attributed slow processing speeds as a reason for their limited use. However, the *Kjernejournalen* software provides a function that was highly favored by all the Norwegian physicians in this study—the pharmaceutical tracking function. The *Kjernejournalen* connects with all the pharmacies in the country and updates automatically as prescriptions are filled [55,56]. This tool was favored as it provides useful information without increasing data entry responsibilities. When asked if the Core Journal has affected their medical practice, the first participant provided the following response:

> *I would say that one way is you can now see what is prescribed and if it has been collected. It is a more secure way of finding out what patients are really taking.* [Norway, participant 1]

Overall, there were mixed feelings about the software among Norwegian physicians. In contrast, all the American physicians expressed their desire for something similar upon learning about the Core Journal. They were also interested in the pharmaceutical tracking function, specifically for narcotic medications. Of the 50 states, all except one (Missouri) have state-wide tracking software; however, communication between programs is limited [86]. In addition, the American cohort at the KUMC faces the unique challenge of working within a facility that is geographically located on the Kansas–Missouri state-line border.

Interoperability

In-depth conversation regarding EHR interoperability capabilities revealed significantly different experiences between the EDs in the 2 countries. In Norway, specialty care is confined to hospitals and allows EM physicians to easily view specialist or inpatient notes. However, primary care facilities are part of the private sector and use different EHRs. Hospital and primary physicians alike are able to access the Core Journal, which provides information regarding critical diagnosis and current

prescriptions [55]. However, Norwegian participants indicated that emergency care was never impeded because of the inability to access primary care clinic notes. Instead, their frustration occurred when requesting imaging from distant facilities. Both cohorts reported needing outside records and imaging occasionally. All the physicians found this task to be annoying and time-consuming. In Norway, all radiologic studies can be electronically exchanged among health care systems throughout the country and sometimes require several phone calls. It was reported that this can take up to 20 minutes but is typically completed more quickly. American physicians noted that they can occasionally view outside imaging. However, this is often not available and scans have to be repeated.

Patients in the United States often receive both primary or specialist care in an outpatient clinic setting. Providers have limited access to patient information at the point of care if health care facilities use different HIT suppliers [12,87]. Over the past decade, laws have been passed with the goal of improving interoperability, but definitive legal parameters are yet to be firmly established [12,45,46]. HIE configuration decisions are typically dependent on the competing vendors and participating health care systems, with both parties having significant effect on user accessibility [48]. Vendors have capitalized on developing exchange capabilities as a product selling point [46]. Subsequently, there have been calls for stronger legislative regulation to improve transparency across health care facilities [41,45,46].

Although individual EHR suppliers have improved interoperability, substantial limitations persist [19]. For example, American physicians in this study discussed the Care Everywhere platform within Epic Systems that grants access to most outpatient documentation and laboratory results from another large hospital within Kansas City. However, this tool still omits numerous facilities and hospitals. US participants reported that electronic exchanges between unaffiliated health care facilities are either impossible or extremely cumbersome and time-consuming. A participant described the process used to request outside medical records and said that it could take hours to days to receive a fax that potentially contains critical information. Knowing that the information will not be available within their own shift, this participant typically makes these requests to benefit colleagues who are taking over patient care. American physicians also believed that redundant diagnostic tests are a direct result of limited interoperability that increases both patient risk and national health care expenses:

> *We repeat so many x-rays, labs, and scans just because we can't see what was done a day ago. There are deficits in care due to poor EHR interoperability. Today, in this emergency department, there will be an issue because they [outside EHRs] don't communicate.* [United States, participant 6]

The VA is a government-run national health care system that internally developed its own EHR software known as Veterans Information Systems and Technology Architecture [41]. Each US participant who mentioned past VA experiences recalled positive experience with this EHR. Although the participants described the software's interface as cumbersome and rudimental, all of them commented about how it allowed them to provide more comprehensive care because of the ability to access all the pertinent information from any VA facility. Despite the recent contract with Cerner, it will likely take longer than 10 years to finalize the implementation of this software as the sole HIT supplier to all VA facilities [41,49].

Legal Considerations

Another burden unique to Americans is the extensive documentation for legal protection. A recent study showed that approximately 51% of EM physicians in the United States will be sued during their career despite appropriate medical management [88]. This was foreign to Norwegians who rely on the Norwegian Medical Association (NMA) for legal counsel and protection [89]. NMA also functions as a professional society and labor union that annually negotiates with the government on behalf of physicians regarding fair working conditions, compensation, and leave-time [89]. Nearly all Norwegian physicians are NMA members, whereas only 11.4% of American health care providers are unionized [90]. Explanations for low involvement include convoluted multi-payer systems, restrictive federal and state laws, and social stigma [90-92].

Defensive medicine is a normalized practice within US medicine. American EM physicians face approximately a 7.5% annual risk of litigation [93]. Consequently, excessive documentation becomes an essential burden to protect oneself from potential legal ramifications. This liability heavily influences medical decision-making, resulting in excessive workups and hospital admissions. A study of 824 physicians in the United States found that 93% of them reported regular practice of defensive medicine [94]. Of those, more than half of the EM physicians reported using computed tomography, magnetic resonance imaging, or radiography that was not clinically necessary [94].

Responses from the American participants correlated with these findings and many believed that improved interoperability between EHR systems could mitigate these practices while simultaneously decreasing physician litigation anxiety. American participants also noted numerous disadvantages associated with defensive medicine on a societal, patient, and health care provider perspectives; however, abandoning this practice puts the physician at an undeniable risk:

> *A lawyer can go through and subpoena every keystroke made from the moment you enter the record, what is done before completing the note, and if you changed anything. We are humans and will make mistakes. If you type something wrong, it can potentially be used against you to criticize your medical judgment. If I have a learner (i.e., scribe or resident) who wrote something wrong and I change or delete it, that may be held against me.* [United States, participant 6]

This is in stark contrast to the practices in Norway, where physicians pay a small percentage of their salaries to a collective pool within the NMA. If a patient is entitled to compensation, it comes from these funds. All Norwegian participants expressed

that this was a fair and equitable process without many disadvantages, and one physician stated the following:

> *I am only concerned for malpractice because I am always concerned with doing the right thing for my patient. I am not concerned about repercussions for making a mistake. When something goes wrong, we are good at protecting each other and focusing on system errors, not personal ones.* [Norway, participant 2]

Limitations and Future Implications

Our study has several limitations. Qualitative research restricts the use of formal statistical analysis as broad-ranging emotions reduce its reproducibility. These challenges were amplified by complex sociopolitical–technological variations. Generalizability is limited owing to the small sample size and single-center analysis in each country. Therefore, we can only extrapolate speculations to explain the results of this study. No definitive conclusions can be made regarding EHR user satisfaction between the 2 countries. Although this study specifically recruited EM physicians, future research may benefit from expanding to other specialties across multiple facilities. Despite the semistructured interviews having reproducibility limitations, this method was necessary to understand the health care infrastructure and nuances of daily practice within each location. New questions emerged as more information was gained. Although this approach creates inconsistencies, it permits flexibility that is otherwise impossible to achieve using alternative qualitative methods such as surveys. These humanistic interactions are both a strength and weakness of semistructured interviews. Objective metrics regarding usability and satisfaction are difficult to produce with countless independent variables. Nevertheless, this comparison provides rich insight.

Numerous potential factors that may contribute to poor EHR user experiences were identified during the first phase of data collection (American interviews). Much of this occurred without prompts from the interviewer (GG). If these factors did not come up organically in Norwegian physician interviews, the interviewer asked targeted questions pertaining to these topics with the intention of identifying similarities or differences. Although this does not alter the United States' findings, it may artificially inflate Norwegian results regarding perceived EHR disadvantages.

A study by Tutty et al [14] described factors that may enhance EHR experiences and suggested that policy makers, software developers, HIT vendors, payers, health administrators, and users alike may be capable of contributing to collective improvements. They also identified administrative tasks that add to documentation burden, including extensive order entries, billing regulations, coding standards, quality improvement reporting, and system security [14]. As Colicchio et al [40] noted, it is important to consider that national EHRs may not provide the desired insight for future informatics research, as local configurations are customizable even when supplied by the same vendor. After the *Helseplattformen* is implemented in Norway, prospective longitudinal studies measuring similar outcomes may produce additional meaningful information. This novel investigation suggests a framework for theoretical EHR optimization on a global scale. Although the results of this study are not entirely generalizable, it provides a foothold for future research and may stimulate innovative HIT advancements. Additional studies that compare international experiences while considering social and political differences are needed to identify the components that most significantly influence user satisfaction.

Conclusions

This qualitative study explores factors that influence EHR user satisfaction among practicing EM physicians in 2 countries. All the participants believed that this technology has increased their workload while simultaneously acknowledging their heavy reliance on it. They agreed that EHRs are here to stay. The results show that both American and Norwegian physicians experience frustration with EHR, but overall, the United States cohort had significantly more complaints. Participant-driven conversations revealed that each country had moderately differing sources of frustration. Norwegian complaints revolved around intrinsic technical issues. Strategies to mitigate these problems are currently underway as evidenced by the *Én Innbygger–Én Journal* and *Helseplattformen* initiative. Americans harshly criticized the *business of medicine* that they felt was manifested in every facet of HIT implementation. These findings enhance the theory that policies and administration may influence usability to a greater degree than technology itself [9,14,24,26,95].

Use of in-depth, semistructured interviews permitted a deeper understanding of both health care systems. This knowledge was subsequently integrated throughout data analysis and interpretation. The development and use of EHRs is influenced by lawmakers, payers, companies, and regulatory entities. Decisions made by those who are not primary users have a profound impact on the practice of those who use this technology daily. Both countries in this study are currently undergoing significant changes. Norway is poised to make a complete national overhaul of their EHR, and the United States is struggling to reform a vast, expensive, and inefficient health system. If HIT is to be optimized on a global scale, the elements highlighted in this study should be considered when establishing policy, strategy, and vision for the future.

Acknowledgments

The authors would like to thank the following individuals for helping them with participant recruitment and overall project support: Bradly Barth, MD; Peyman Usefi, MD; Marius Aagaard, MD; and Morgan Denton. The authors are grateful to all the participating physicians at the University of Kansas (KU) Health System and Akershus Universitetssykehus. They would also like to acknowledge Hanah Valeur, MD and Eirik Nikolai Arnesen, MD for providing insight into the Norwegian Medical Association and the structure

of the Norwegian health care system. This research was supported by the Clendening and King Summer Fellowship through the KU School of Medicine–Department of History and Philosophy of Medicine. The processing charges related to the publication of this paper were supported by the KU One University Open Access Author Fund sponsored jointly by the KU Provost, KU Vice Chancellor for Research and Graduate Studies, and the KU Medical Center Vice Chancellor for Research and managed jointly by the Libraries at the Medical Center and KU-Lawrence.

Conflicts of Interest

None declared.

References

1. Gardner RL, Cooper E, Haskell J, Harris DA, Poplau S, Kroth PJ, et al. Physician stress and burnout: the impact of health information technology. J Am Med Inform Assoc 2019 Feb 01;26(2):106-114. [doi: 10.1093/jamia/ocy145] [Medline: 30517663]
2. Denton CA, Soni HC, Kannampallil TG, Serrichio A, Shapiro JS, Traub SJ, et al. Emergency physicians' perceived influence of ehr use on clinical workflow and performance metrics. Appl Clin Inform 2018 Jul;9(3):725-733 [FREE Full text] [doi: 10.1055/s-0038-1668553] [Medline: 30208497]
3. Perry JJ, Sutherland J, Symington C, Dorland K, Mansour M, Stiell IG. Assessment of the impact on time to complete medical record using an electronic medical record versus a paper record on emergency department patients: a study. Emerg Med J 2014 Dec;31(12):980-985. [doi: 10.1136/emermed-2013-202479] [Medline: 23975593]
4. Hill RG, Sears LM, Melanson SW. 4000 clicks: a productivity analysis of electronic medical records in a community hospital ED. Am J Emerg Med 2013 Nov;31(11):1591-1594. [doi: 10.1016/j.ajem.2013.06.028] [Medline: 24060331]
5. Neri PM, Redden L, Poole S, Pozner CN, Horsky J, Raja AS, et al. Emergency medicine resident physicians' perceptions of electronic documentation and workflow: a mixed methods study. Appl Clin Inform 2015;6(1):27-41 [FREE Full text] [doi: 10.4338/ACI-2014-08-RA-0065] [Medline: 25848411]
6. Park SY, Lee SY, Chen Y. The effects of EMR deployment on doctors' work practices: a qualitative study in the emergency department of a teaching hospital. Int J Med Inform 2012 Mar;81(3):204-217. [doi: 10.1016/j.ijmedinf.2011.12.001] [Medline: 22217802]
7. Hecht J. The future of electronic health records. Nature 2019 Sep;573(7775):114-116. [doi: 10.1038/d41586-019-02876-y] [Medline: 31554991]
8. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan W, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using ehr event log data and time-motion observations. Ann Fam Med 2017 Sep;15(5):419-426 [FREE Full text] [doi: 10.1370/afm.2121] [Medline: 28893811]
9. Friedberg MW, Chen PG, Van Busum KR, Aunon F, Pham C, Caloyeras J, et al. Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. Rand Health Q 2014;3(4):1 [FREE Full text] [Medline: 28083306]
10. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med 2016 Sep 6:753-760. [doi: 10.7326/M16-0961] [Medline: 27595430]
11. Khairat S, Coleman C, Newlin T, Rand V, Ottmar P, Bice T, et al. A mixed-methods evaluation framework for electronic health records usability studies. J Biomed Inform 2019 Jun;94:103175 [FREE Full text] [doi: 10.1016/j.jbi.2019.103175] [Medline: 30981897]
12. Office of the National Coordinator for Health Information Technology (ONC). 2018 Report to Congress. Department of Health & Human Services (HHS). 2018. URL: https://www.healthit.gov/sites/default/files/page/2018-12/2018-HITECH-report-to-congress.pdf [accessed 2020-12-01]
13. Hertzum M, Ellingsen G. The implementation of an electronic health record: comparing preparations for epic in Norway with experiences from the UK and Denmark. Int J Med Inform 2019 Sep;129:312-317. [doi: 10.1016/j.ijmedinf.2019.06.026] [Medline: 31445272]
14. Tutty MA, Carlasare LE, Lloyd S, Sinsky CA. The complex case of EHRs: examining the factors impacting the EHR user experience. J Am Med Inform Assoc 2019 Jul 01;26(7):673-677 [FREE Full text] [doi: 10.1093/jamia/ocz021] [Medline: 30938754]
15. Garavand A, Mohseni M, Asadi H, Etemadi M, Moradi-Joo M, Moosavi A. Factors influencing the adoption of health information technologies: a systematic review. Electron Physician 2016 Aug;8(8):2713-2718 [FREE Full text] [doi: 10.19082/2713] [Medline: 27757179]
16. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2016 May 20;Suppl 1:48-61. [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
17. Kaipio J, Lääveri T, Hyppönen H, Vainiomäki S, Reponen J, Kushniruk A, et al. Usability problems do not heal by themselves: National Survey on physicians' experiences with EHRs in Finland. Int J Med Inform 2017 Jan;97:266-281 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.10.010] [Medline: 27919385]

18. Jha AK, Doolan D, Grandt D, Scott T, Bates DW. The use of health information technology in seven nations. Int J Med Inform 2008 Dec;77(12):848-854. [doi: 10.1016/j.ijmedinf.2008.06.007] [Medline: 18657471]
19. Payne TH, Lovis C, Gutteridge C, Pagliari C, Natarajan S, Yong C, et al. Status of health information exchange: a comparison of six countries. J Glob Health 2019 Dec;9(2):0204279 [FREE Full text] [doi: 10.7189/jogh.09.020427] [Medline: 31673351]
20. Teixeira JG, Pinho NF, Patrício L. Bringing service design to the development of health information systems: the case of the Portuguese national electronic health record. Int J Med Inform 2019 Dec;132:103942. [doi: 10.1016/j.ijmedinf.2019.08.002] [Medline: 31627031]
21. Powell AC, Ludhar JK, Ostrovsky Y. Electronic health record use in an affluent region in India: findings from a survey of Chandigarh hospitals. Int J Med Inform 2017 Dec;103:78-82. [doi: 10.1016/j.ijmedinf.2017.04.011] [Medline: 28551005]
22. Salleh MI, Abdullah R, Zakaria N. Evaluating the effects of electronic health records system adoption on the performance of Malaysian health care providers. BMC Med Inform Decis Mak 2021 Feb 25;21(1):75 [FREE Full text] [doi: 10.1186/s12911-021-01447-4] [Medline: 33632216]
23. Kushniruk AW, Bates DW, Bainbridge M, Househ MS, Borycki EM. National efforts to improve health information system safety in Canada, the United States of America and England. Int J Med Inform 2013 May;82(5):149-160. [doi: 10.1016/j.ijmedinf.2012.12.006] [Medline: 23313431]
24. Downing NL, Bates DW, Longhurst CA. Physician burnout in the electronic health record era: are we ignoring the real cause? Ann Intern Med 2018 Jul 03;169(1):50-51. [doi: 10.7326/M18-0139] [Medline: 29801050]
25. Faber S, van Geenhuizen M, de Reuver M. eHealth adoption factors in medical hospitals: a focus on the Netherlands. Int J Med Inform 2017 Apr;100:77-89. [doi: 10.1016/j.ijmedinf.2017.01.009] [Medline: 28241940]
26. Whyle E, Olivier J. Social values and health systems in health policy and systems research: a mixed-method systematic review and evidence map. Health Policy Plan 2020 Jul 01;35(6):735-751 [FREE Full text] [doi: 10.1093/heapol/czaa038] [Medline: 32374881]
27. Halamka JD, Tripathi M. The HITECH era in retrospect. N Engl J Med 2017 Sep 07;377(10):907-909. [doi: 10.1056/NEJMp1709851] [Medline: 28877012]
28. Sawyer NT. Why the EMTALA mandate for emergency care does not equal healthcare "Coverage". West J Emerg Med 2017 Jun;18(4):551-552 [FREE Full text] [doi: 10.5811/westjem.2017.5.34826] [Medline: 28611872]
29. Zibulewsky J. The Emergency Medical Treatment and Active Labor Act (EMTALA): what it is and what it means for physicians. Proc (Bayl Univ Med Cent) 2001 Oct;14(4):339-346 [FREE Full text] [doi: 10.1080/08998280.2001.11927785] [Medline: 16369643]
30. EMTALA fact sheet. American College of Emergency Physicians. URL: https://www.acep.org/life-as-a-physician/ethics--legal/emtala/emtala-fact-sheet/ [accessed 2020-06-01]
31. Shenvi C, Beise K, Tintinalli J. 30 años de programas de residencia en Medicina de Urgencias y Emergencias en Estados Unidos. Emergencias. 2013. URL: https://www.usfx.bo/nueva/vicerrectorado/citas/SALUD_10/Enfermeria_Obtetriz/98.pdf [accessed 2020-10-01]
32. Augustine J. Latest data reveal the ED's role as hospital admission gatekeeper. American College of Emergency Physicians. 2019. URL: https://www.acepnow.com/article/latest-data-reveal-the-eds-role-as-hospital-admission-gatekeeper/ [accessed 2020-12-01]
33. Gindi RM, Black LI, Cohen RA. Reasons for Emergency Room Use Among U.S. Adults Aged 18-64: National Health Interview Survey, 2013 and 2014. Natl Health Stat Report 2016 Feb 18(90):1-16 [FREE Full text] [Medline: 26905514]
34. Consumer expenditure survey - Health insurance. U.S. Bureau of Labor Statistics. 2020. URL: https://www.bls.gov/cex/ [accessed 2021-06-01]
35. Schoen C, Osborn R, Squires D, Doty MM, Pierson R, Applebaum S. How health insurance design affects access to care and costs, by income, in eleven countries. Health Aff (Millwood) 2010 Dec;29(12):2323-2334. [doi: 10.1377/hlthaff.2010.0862] [Medline: 21088012]
36. Capp R, Rooks SP, Wiler JL, Zane RD, Ginde AA. National study of health insurance type and reasons for emergency department use. J Gen Intern Med 2014 Apr;29(4):621-627 [FREE Full text] [doi: 10.1007/s11606-013-2734-4] [Medline: 24366398]
37. Richman IB, Clark S, Sullivan AF, Camargo CA. National study of the relation of primary care shortages to emergency department utilization. Acad Emerg Med 2007 Mar;14(3):279-282 [FREE Full text] [doi: 10.1197/j.aem.2006.10.093] [Medline: 17242383]
38. Washington V, DeSalvo K, Mostashari F, Blumenthal D. The HITECH era and the path forward. N Engl J Med 2017 Sep 07;377(10):904-906. [doi: 10.1056/NEJMp1703370] [Medline: 28877013]
39. Meaningful use: Electronic Health Record (EHR) incentive programs. American Medical Association. 2020. URL: https://www.ama-assn.org/practice-management/medicare/meaningful-use-electronic-health-record-ehr-incentive-programs [accessed 2020-10-01]
40. Colicchio TK, Cimino JJ, Del Fiol G. Unintended consequences of nationwide electronic health record adoption: challenges and opportunities in the post-meaningful use era. J Med Internet Res 2019 Jun 03;21(6):e13313 [FREE Full text] [doi: 10.2196/13313] [Medline: 31162125]

41. Reisman M. EHRs: the challenge of making electronic data usable and interoperable. P T 2017 Sep;42(9):572-575 [FREE Full text] [Medline: 28890644]
42. Ratwani RM, Fairbanks RJ, Hettinger AZ, Benda NC. Electronic health record usability: analysis of the user-centered design processes of eleven electronic health record vendors. J Am Med Inform Assoc 2015 Nov 01;22(6):1179-1182 [FREE Full text] [doi: 10.1093/jamia/ocv050] [Medline: 26049532]
43. Schulte F, Fry E. Death by 1,000 clicks: where electronic health records went wrong. Kaiser Health News. 2019. URL: https://khn.org/news/death-by-a-thousand-clicks/ [accessed 2019-10-20]
44. Non-federal acute care hospital electronic health record adoption. Health IT Dashboard. URL: https://dashboard.healthit.gov/quickstats/pages/FIG-Hospital-EHR-Adoption.php [accessed 2020-11-25]
45. Black JR, Hulkower RL, Ramanathan T. Health Information Blocking: Responses Under the 21st Century Cures Act. Public Health Rep 2018;133(5):610-613 [FREE Full text] [doi: 10.1177/0033354918791544] [Medline: 30134128]
46. Adler-Milstein J, Pfeifer E. Information blocking: is it occurring and what policy strategies can address it? Milbank Q 2017 Mar 07;95(1):117-135 [FREE Full text] [doi: 10.1111/1468-0009.12247] [Medline: 28266065]
47. Health information technology: certification and interoperability enhancements. Department of Health and Human. 2018. URL: https://www.reginfo.gov/public/do/eAgendaViewRule?pubId=201804&RIN=0955-AA01 [accessed 2021-01-18]
48. Downing NL, Adler-Milstein J, Palma JP, Lane S, Eisenberg M, Sharp C, Northern California HIE Collaborative, et al. Health information exchange policies of 11 diverse health systems and the associated impact on volume of exchange. J Am Med Inform Assoc 2017 Jan;24(1):113-122 [FREE Full text] [doi: 10.1093/jamia/ocw063] [Medline: 27301748]
49. Cerner, VA to provide Veterans with seamless care. Federal Update. 2018. URL: https://www.cerner.com/perspectives/cerner-va-to-provide-veterans-with-seamless-care [accessed 2020-11-28]
50. Ringard AS, Sagan A, Sperre Saunes I, Lindahl AK. Norway: health system review. Health Syst Transit 2013;15(8):1-162 [FREE Full text] [Medline: 24434287]
51. What is core journal? Norwegian Health Network. 2020. URL: https://www.helsenorge.no/kjernejournal/hva-er-kjernejournal/#forskjellen-mellom-pasientjournal-og-kjernejournal [accessed 2021-08-09]
52. Galletta G, Løvstakken K. Emergency medicine in Norway: the road to specialty recognition. J Am Coll Emerg Physicians Open 2020 Oct;1(5):790-794 [FREE Full text] [doi: 10.1002/emp2.12197] [Medline: 33145520]
53. Mentzoni I, Bogstrand ST, Faiz KW. Emergency department crowding and length of stay before and after an increased catchment area. BMC Health Serv Res 2019 Jul 22;19(1):506 [FREE Full text] [doi: 10.1186/s12913-019-4342-4] [Medline: 31331341]
54. Arnesen EN. E-Health in Norway. E-health in Hospital Care in 2017. Luxembourg: European Health Management Association (EHMA) Conference; 2017 May. URL: https://www.aemh.org/images/AEMH_documents/2017/Arnesen.pdf [accessed 2019-06-28]
55. Arnesen EN, Larsen BA. Alert information in the Norwegian Summary Care Record. Tidsskr Nor Laegeforen 2014 Oct 28;134(20):1927-1928 [FREE Full text] [doi: 10.4045/tidsskr.14.1085] [Medline: 25350434]
56. Dyb K, Warth LL. The Norwegian National Summary Care Record: a qualitative analysis of doctors' use of and trust in shared patient information. BMC Health Serv Res 2018 Apr 06;18(1):252 [FREE Full text] [doi: 10.1186/s12913-018-3069-y] [Medline: 29625587]
57. Warth L, Dyb K. A qualitative study of the implementation and use of a national information system. International Association for Development of the Information Society (IADIS). 2017. URL: https://www.researchgate.net/publication/320622525_A_QUALITATIVE_STUDY_OF_THE_IMPLEMENTATION_AND_USE_OF_A_NATIONAL_INFORMATION_SYSTEM [accessed 2021-01-18]
58. One inhabitant- one journal. Ministry of Health and Care Services. 2013. URL: https://www.regjeringen.no/no/dokumenter/meld-st-9-20122013/id708609/?ch=1 [accessed 2021-01-18]
59. Green A. EMR/EHR developments in the Nordics - White paper. Signify Research Knowledge Centre. 2018. URL: https://s3-eu-west-2.amazonaws.com/signifyresearch/app/uploads/2018/12/17103930/EHR-EMR-Market-Developments-in-the-Nordics-White-Paper-17-Dec-20181.pdf [accessed 2021-01-17]
60. Onarheim K. Helseplattformen- one health record across all levels of care- a ground-breaking project in Central Norway. In: Proceedings of the European Conference and Exhibition. 2019 Jun Presented at: European Conference and Exhibition; June 11-13, 2019; Helsinki, Finland URL: https://www.himss.eu/sites/himsseu/files/site-assets/community/nordic-19/Kristian%20Onarheim_One%20citizen-one%20journal%20in%20Norway.pdf
61. Helseplattformen AS. Helseplattformen. 2019. URL: https://helseplattformen.no/om-oss/helseplattformen-as [accessed 2020-11-21]
62. Melby LM, Andreassen HK, Torsvik T, Ellingsen G, Severinsen GH, Silsand L, et al. Ambivalently awaiting: Norwegian general practitioners' expectations towards a cross-institutional electronic health record. In: Proceedings of the Linköping Electronic Conference. 2019 Presented at: Linköping Electronic Conference; 2019; Oslo, Norway p. 99-103 URL: https://munin.uit.no/handle/10037/18064
63. The project. Helseplattformen. 2020. URL: https://helseplattformen.no/om-oss/prosjektet [accessed 2020-11-27]
64. Effort-driven financing (ISF) - regulations. Helsedirektoratet. 2020. URL: https://www.helsedirektoratet.no/tema/finansiering/innsatsstyrt-finansiering-og-drg-systemet/innsatsstyrt-finansiering-isf [accessed 2020-11-22]

65. Tikkanen R, Osborn R, Mossialos E, Djordjevic A, Wharton GA. Norway. The Commonwealth Fund. URL: https://www.commonwealthfund.org/international-health-policy-center/countries/norway [accessed 2019-09-03]
66. Blome A, Yu D, Lu X, Schreyer KE. Pitfalls of Extensive Documentation in the Emergency Department. Ochsner J 2020 Sep 25;20(3):299-302 [FREE Full text] [doi: 10.31486/toj.19.0108] [Medline: 33071663]
67. International Classification of Diseases Eleventh Revision (ICD-11). National Committee on Vital and Health Statistics (NCVHS).: U.S. Department of Health and Human Services; 2019 Oct 10. URL: https://ncvhs.hhs.gov/wp-content/uploads/2020/02/Final-Meeting-Summary-Standards-Subcommittee-ICD-11-Meeting.pdf [accessed 2020-11-24]
68. Levinson DR. Improper payments for evaluation and management services cost medicare billions in 2010. Department of Health and Human Services. 2014. URL: https://oig.hhs.gov/oei/reports/oei-04-10-00181.pdf [accessed 2020-11-20]
69. Proctor J. Gauging emergency physician productivity: are RVUs the answer? American College of Emergency Physicians. URL: https://www.acep.org/imports/clinical-and-practice-management/resources/reimbursement-imported/gauging-emergency-physician-productivity-are-rvus-the-answer/ [accessed 2021-01-18]
70. Gottlieb M, Palter J, Westrick J, Peksa GD. Effect of medical scribes on throughput, revenue, and patient and provider satisfaction: a systematic review and meta-analysis. Ann Emerg Med 2021 Feb;77(2):180-189. [doi: 10.1016/j.annemergmed.2020.07.031] [Medline: 32868143]
71. Tajirian T, Stergiopoulos V, Strudwick G, Sequeira L, Sanches M, Kemp J, et al. The influence of electronic health record use on physician burnout: cross-sectional survey. J Med Internet Res 2020 Jul 15;22(7):e19274 [FREE Full text] [doi: 10.2196/19274] [Medline: 32673234]
72. Khairat S, Burke G, Archambault H, Schwartz T, Larson J, Ratwani R. Perceived burden of EHRs on physicians at different stages of their career. Appl Clin Inform 2018 Apr 16;9(2):336-347 [FREE Full text] [doi: 10.1055/s-0038-1648222] [Medline: 29768634]
73. Robertson SL, Robinson MD, Reid A. Electronic Health Record Effects on Work-Life Balance and Burnout Within the I Population Collaborative. J Grad Med Educ 2017 Aug;9(4):479-484 [FREE Full text] [doi: 10.4300/JGME-D-16-00123.1] [Medline: 28824762]
74. Shanafelt TD, Dyrbye LN, Sinsky C, Hasan O, Satele D, Sloan J, et al. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. Mayo Clin Proc 2016 Jul;91(7):836-848. [doi: 10.1016/j.mayocp.2016.05.007] [Medline: 27313121]
75. Shanafelt TD, Noseworthy JH. Executive leadership and physician well-being: nine organizational strategies to promote engagement and reduce burnout. Mayo Clin Proc 2016 Nov 18:129-146. [doi: 10.1016/j.mayocp.2016.10.004] [Medline: 27871627]
76. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med 2018 Jun;283(6):516-529 [FREE Full text] [doi: 10.1111/joim.12752] [Medline: 29505159]
77. How doctors feel about electronic health records: the Harris poll. Stanford Medicine. 2018. URL: https://med.stanford.edu/content/dam/sm/ehr/documents/EHR-Poll-Presentation.pdf [accessed 2021-01-17]
78. Mahmood JI, Grotmol KS, Tesli M, Moum T, Andreassen O, Tyssen R. Life satisfaction in Norwegian medical doctors: a 15-year longitudinal study of work-related predictors. BMC Health Serv Res 2019 Oct 22;19(1):729 [FREE Full text] [doi: 10.1186/s12913-019-4599-7] [Medline: 31640717]
79. Hertzberg TK, Skirbekk H, Tyssen R, Aasland OG, Rø KI. The hospital doctor of today - still continuously on duty. Tidsskr Nor Laegeforen 2016 Oct;136(19):1635-1638 [FREE Full text] [doi: 10.4045/tidsskr.16.0067] [Medline: 27790890]
80. Røvik JO, Tyssen R, Hem E, Gude T, Ekeberg O, Moum T, et al. Job stress in young physicians with an emphasis on the work-home interface: a nine-year, nationwide and longitudinal study of its course and predictors. Ind Health 2007 Oct;45(5):662-671 [FREE Full text] [doi: 10.2486/indhealth.45.662] [Medline: 18057809]
81. Tyssen R. Health problems and the use of health services among physicians: a review article with particular emphasis on Norwegian studies. Ind Health 2007 Oct;45(5):599-610 [FREE Full text] [doi: 10.2486/indhealth.45.599] [Medline: 18057803]
82. Rø KEI, Gude T, Aasland OG. Does a self-referral counselling program reach doctors in need of help? A comparison with the general Norwegian doctor workforce. BMC Public Health 2007 Mar 16;7:36 [FREE Full text] [doi: 10.1186/1471-2458-7-36] [Medline: 17367526]
83. Isaksson Ro KE, Tyssen R, Gude T, Aasland OG. Will sick leave after a counselling intervention prevent later burnout? A 3-year follow-up study of Norwegian doctors. Scand J Public Health 2012 May;40(3):278-285 [FREE Full text] [doi: 10.1177/1403494812443607] [Medline: 22637367]
84. Langballe EM, Innstrand ST, Hagtvet KA, Falkum E, Gjerløw Aasland O. The relationship between burnout and musculoskeletal pain in seven Norwegian occupational groups. Work 2009;32(2):179-188. [doi: 10.3233/WOR-2009-0804] [Medline: 19289871]
85. Stehman CR, Testo Z, Gershaw RS, Kellogg AR. Burnout, drop out, suicide: physician loss in emergency medicine, part I. West J Emerg Med 2019 May;20(3):485-494 [FREE Full text] [doi: 10.5811/westjem.2019.4.40970] [Medline: 31123550]
86. Weber L. Why Missouri's the Last Holdout on a Statewide Rx Monitoring Program. Kaiser Health News. 2019. URL: https://khn.org/news/why-missouris-the-last-holdout-on-a-statewide-rx-monitoring-program/ [accessed 2020-11-29]

87. Kern LM, Seirup JK, Rajan M, Jawahar R, Stuard SS. Fragmented ambulatory care and subsequent healthcare utilization among Medicare beneficiaries. Am J Manag Care 2018 Sep 01;24(9):278-284 [FREE Full text] [Medline: 30222925]
88. Guardado JR. Medical liability claim frequency among U.S. physicians. American Medical Association. 2017. URL: https://www.ama-assn.org/sites/ama-assn.org/files/corp/media-browser/public/government/advocacy/policy-research-perspective-medical-liability-claim-frequency.pdf [accessed 2021-01-17]
89. The role of the Norwegian Medical Association. Den Norske Legeforening. 2018. URL: https://www.legeforeningen.no/om-oss/english/ [accessed 2021-01-17]
90. Howard D. What should physicians consider prior to unionizing? AMA J Ethics 2020 Mar 01;22(3):193-200 [FREE Full text] [doi: 10.1001/amajethics.2020.193] [Medline: 32220265]
91. Adams MA, Allen JI. Understanding the Legal and Regulatory Framework Governing Physician Collective Bargaining. Am J Gastroenterol 2018 Feb;113(2):170-171. [doi: 10.1038/ajg.2017.424] [Medline: 29160299]
92. Choudhry S, Brennan TA. Collective bargaining by physicians--labor law, antitrust law, and organized medicine. N Engl J Med 2001 Oct 11;345(15):1141-1144. [doi: 10.1056/NEJM200110113451522] [Medline: 11596605]
93. Ferguson B, Geralds J, Petrey J, Huecker M. Malpractice in emergency medicine-a review of risk and mitigation practices for the emergency medicine provider. J Emerg Med 2018 Nov;55(5):659-665. [doi: 10.1016/j.jemermed.2018.06.035] [Medline: 30166074]
94. Studdert DM, Mello MM, Sage WM, DesRoches CM, Peugh J, Zapert K, et al. Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. J Am Med Assoc 2005 Jun 01;293(21):2609-2617. [doi: 10.1001/jama.293.21.2609] [Medline: 15928282]
95. 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):2-8 [FREE Full text] [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]

Abbreviations

ED: emergency department
EHR: electronic health record
EM: emergency medicine
HIE: health information exchange
HIT: health information technology
KU: University of Kansas
KUMC: University of Kansas Medical Center
NMA: Norwegian Medical Association
VA: US Department of Veterans Affairs

Edited by A Kushniruk; submitted 14.03.21; peer-reviewed by L Kremer, A Follmann; comments to author 24.05.21; revised version received 22.08.21; accepted 11.10.21; published 07.01.22.

Please cite as:
Garcia G, Crenner C
Comparing International Experiences With Electronic Health Records Among Emergency Medicine Physicians in the United States and Norway: Semistructured Interview Study
JMIR Hum Factors 2022;9(1):e28762
URL: https://humanfactors.jmir.org/2022/1/e28762
doi:10.2196/28762
PMID:34994702

Original Paper

Uses of Personal Health Records for Communication Among Colorectal Cancer Survivors, Caregivers, and Providers: Interview and Observational Study in a Human-Computer Interaction Laboratory

David A Haggstrom[1,2,3], MD, MAS; Thomas Carr[1], MD

[1]VA HSR&D Center for Health Information and Communication, Roudebush Veterans Affairs Medical Center, Indianapolis, IN, United States

[2]Division of General Internal Medicine & Geriatrics, Indiana University School of Medicine, Indianapolis, IN, United States

[3]Center for Health Services Research, Regenstrief Institute, Indianapolis, IN, United States

Corresponding Author:
David A Haggstrom, MD, MAS
VA HSR&D Center for Health Information and Communication
Roudebush Veterans Affairs Medical Center
1481 West 10th Street
Indianapolis, IN, 46202
United States
Phone: 1 317 274 9304
Email: dahaggst@iupui.edu

Abstract

Background: Personal health records (PHRs) may be useful for patient self-management and participation in communication with their caregivers and health care providers. As each potential participant's role is different, their perception of the best uses of a PHR may vary.

Objective: The perspectives of patients, caregivers, and providers were all evaluated concurrently in relation to a PHR developed for colorectal cancer (CRC) survivors.

Methods: We explored group perceptions of a CRC PHR prototype. Scenario-based testing across eight use cases, with semistructured follow-up interviews, was videotaped in a human-computer interaction laboratory with patients, caregivers, and health care providers. Providers included oncologists, gastroenterologists, and primary care physicians. Discrete observations underwent grounded theory visual affinity analysis to identify emergent themes.

Results: Observations fell into three major themes: the network (who should be granted access to the PHR by the patient), functions (helpful activities the PHR enabled), and implementation (how to adopt the PHR into workflow). Patients wanted physician access to their PHR, as well as family member access, especially when they lived at a distance. All groups noted the added value of linking the PHR to an electronic health record, self-tracking, self-management, and secure messaging. Patients and caregivers also saw information in the PHR as a useful memory tool given their visits to multiple doctors. Providers had reservations about patients viewing raw data, which they were not prepared to interpret or might be inaccurate; patients and caregivers did not express any reservations about having access to more information. Patients saw PHR communication functions as a potential tool for relationship building. Patients and caregivers valued the journal as a tool for reflection and delivery of emotional support. Providers felt the PHR would facilitate patient-physician communication but worried that sharing journal access would make the doctor-patient relationship less professional and had reservations about the time burden of reviewing. Strategies suggested for efficient adoption into workflow included team delegation. Establishment of parameters for patient uses and provider responses was perceived as good standard practice.

Conclusions: PHR perceptions differed by role, with providers seeing the PHR as informational, while patients and caregivers viewed the tool as more relational. Personal health records should be linked to electronic health records for ease of use. Tailoring access, content, and implementation of the PHR is essential. Technology changes have the potential to change the nature of the patient-physician relationship. Patients and providers should establish shared expectations about the optimal use of the PHR and explore how emerging patient-centered technologies can be successfully implemented in modern medical practice to improve the relational quality of care.

(JMIR Hum Factors 2022;9(1):e16447) doi:10.2196/16447

KEYWORDS

personal health record; communication; cancer survivorship; colorectal cancer

Introduction

Personal Health Records (PHRs) have grown in popularity and functionality over time. PHRs are grounded in the idea that patients can improve continuity of care by transporting copies of their records from doctor to doctor. Originally, the term "PHR" stood for "patient-held records" and were paper-based systems [1]. Initially designed by patients, institutions later developed standardized PHR formats to include medication lists, test reports, and physician notes [1-3]. The patient empowerment movement and the internet transformed PHRs into "personal health records" [1,4,5]. Advances in information technology have provided new tools for web-based self-management, communication, and information-sharing to enable patients to play a more active role in their care.

In the United States, patients with chronic disease represented the first target populations for PHRs [4,5]; while their adoption rate is only slightly higher than that in the general population, patients with chronic diseases make greater use of PHR capabilities [6]. Many PHR self-management tools are disease-specific, suggesting that the ideal composition might vary by disease [7-9]; this insight has led to the development of specialized PHRs designed for different chronic diseases, such as diabetes, heart disease, or neurologic disorders [10]. With increases in cancer survival, oncology care has increasingly assumed the characteristics of chronic disease management [11]. To meet the needs of long-term cancer survivorship, the Institute of Medicine (IOM) recommended the development of survivorship care plans to provide a treatment summary and plan for follow-up care, including the potential side effects and long-term consequences of treatment, the timing and content of recommended follow-up, and psychosocial services available in the community [12]. PHRs provide a natural platform on which to address these goals; thus, they have been developed for several cancers [5]. Colorectal cancer (CRC) is the second-most common cause of cancer death in the United States, approximately 147,950 individuals will be diagnosed with CRC and 53,200 will die from the disease in 2020 [13]. All of the general issues addressed by survivorship care plans are of specific relevance to CRC survivors, including follow-up care or surveillance (colonoscopy, carcinoembryonic testing, and abdominal imaging [14]) and potential side effects (eg, radiation proctitis [15] or oxaliplatin neuropathy [16]).

The patient is arguably the key stakeholder, or user, in PHR design. Nonetheless, the proliferation of both synchronous and asynchronous methods of communication has expanded the scope of stakeholders to include both caregivers and health care providers. To create the most effective PHR, the design process needs to account for the needs of all potential stakeholders. Prior studies have examined the perspectives of each stakeholder group individually or focused upon combinations of providers [17] or patients and caregivers [18], or even provider perspectives of caregiver use [19]. Nonetheless, the input of stakeholders such as patients, caregivers, and providers are not commonly considered. Prior qualitative research has shown that patients perceive web-based chronic disease management portals as increasing their access to information and engagement in health care, but improvements in portal design may improve usability and reduce attrition. Caregivers have expressed high interest in portal use to support their roles in interpreting health information, advocating for quality care, and managing medical care [18]. Providers have previously described secure messaging as having particular value for both themselves and their patients; however, providers also expressed concern about the inability of patients to share other types of information with their health care team [17] and the impact on workflow. In this study, the perspectives of patients, caregivers, and health care providers were all evaluated concurrently in relation to a PHR developed for CRC survivors. Patient and caregiver engagement is important for the adoption of PHRs, whereas provider buy-in is critical to the implementation of these technologies in health care settings.

Our key study question was what are the areas of agreement and disagreement among patients, caregivers, and providers with respect to the benefits and appropriate uses of a PHR. Across stakeholders, we explored several questions, including "Who should be provided access to, share information, and communicate with the PHR?" and "What type of patient-generated information should be incorporated into the PHR?" Finally, we asked how the PHR impacts workflow and what best practices may guide the future design and implementation of PHRs for patients with cancer.

Methods

Participants

Four to six participants were recruited from each role group (patient, provider, and caregiver) on the basis of a previous study by Nielsen et al [20], suggesting that this number is sufficient to detect the majority of usability problems. Six CRC survivors were recruited from the Roudebush Veterans Affairs Medical Center (RVAMC) oncology clinic in Indianapolis. Provider schedules were reviewed prior to their clinic visit, and then research assistants approached patients in-person at their planned clinic visit. Either during the clinic encounter or later when scheduling the testing session over the telephone, patients were invited to identify a caregiver who could also participate in the session. For inclusion, cancer survivors were required to have a diagnosis of colorectal cancer more than 12 months prior to enrollment. This yearlong interval was chosen to identify patients who were likely to have undergone both surgical and adjuvant therapy so that they could provide feedback on both treatment modalities, as well as to minimize respondent burden upon any patient undergoing active treatment. One caregiver, identified as a family member or friend supporting the cancer survivor's health needs, was recruited along with each cancer survivor. Seven health care providers were purposefully

recruited via email from the RVAMC, including an oncologist, oncology nurse, gastroenterologist, and 4 primary care physicians. In terms of incentives, gift cards were offered in the amount of US $5 to providers and US $25 to patients and caregivers. All participants provided informed consent, and the study protocol was approved by the institutional review board of Indiana University.

Prototype Design

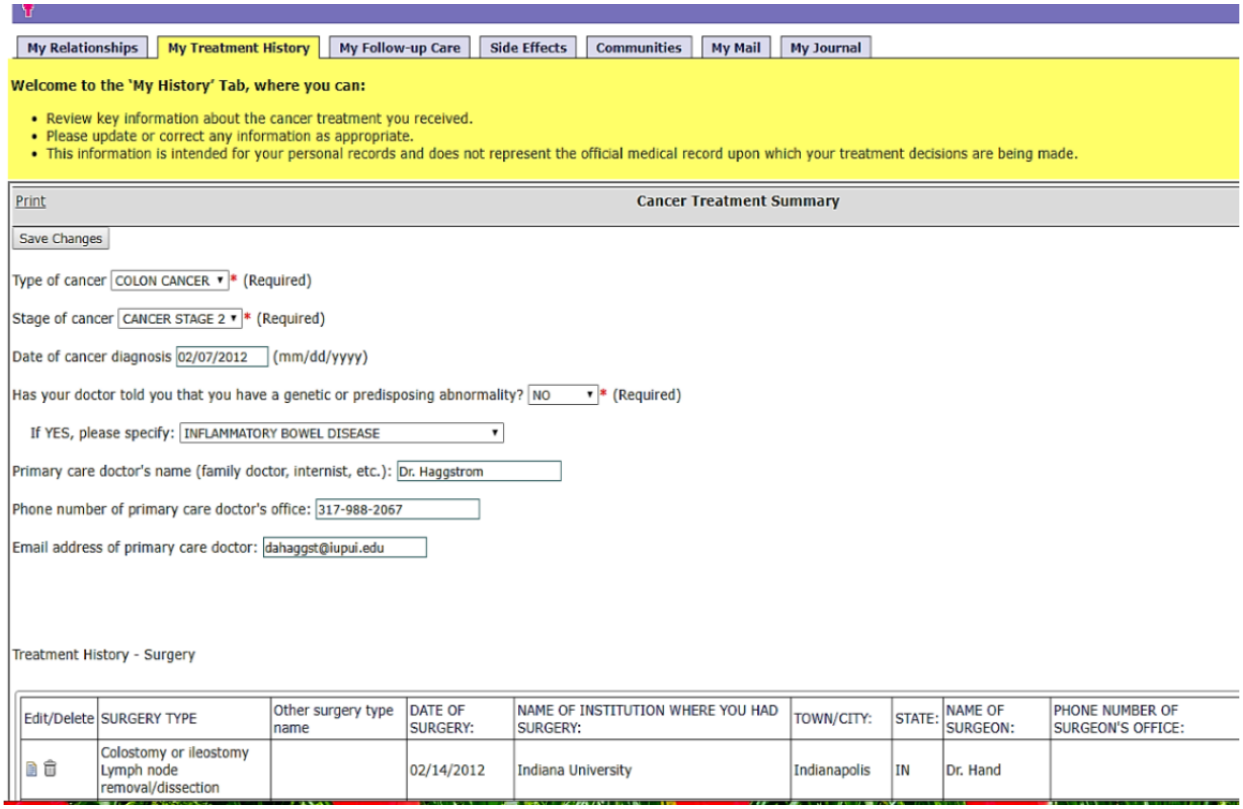
The initial, web-based CRC PHR prototype was created by a design team of clinical investigators and software developers. The functions of this prototype were informed by the IOM report on Cancer Survivorship [12] and are enumerated in Table 1.

The prototype used a tabbed browser format created with open-source software, the OpenMRS medical record system platform [21] shown in a screenshot in Figure 1.

Table 1. Functions of the personal health record of colorectal cancer survivors.

Tab	Functions
My History	Allows review of cancer diagnosis and treatment, including specific type of surgery and adjuvant therapy (chemotherapy, and radiotherapy)
My Follow-up Care	Two tables: a table with recommended surveillance tests based on initial diagnosis and a table of actual tests performed (date, test, and result)
Side Effects	Tailored compendium of possible side effects of treatment and initial, straightforward self-management steps
Communities	Web-based links to cancer information resources and cancer survivor support groups
Relationships	Patients can share access to their personal health records with a set of role-based individuals (provider, caregiver, etc): relationship function enables tiered access to personal (My Mail and My Journal), medical (My History, My Follow-up Care, and Side Effects) or all components of their personal health record
My Mail	Client-based email application enabling secure message exchange
My Journal	Searchable, dated electronic journal, with an ability for in-line comments or responses by individuals to whom a Relationship has been granted

Figure 1. Prototype screenshot created with open-source software—the OpenMRS medical record system platform.



Interviews and Observations

Using the final PHR version, the participants all completed an individual session in the RVAMC human-computer interaction laboratory. Each session lasted 1 hour and was conducted by a member of the study team with a background in human factors engineering. None of the participants were familiar with the PHR prior to enrollment, nor was any type of tutorial provided before use so as to avoid bias on the basis of user experience, as well as assess the usability or intuitiveness of the interface.

An interview guide was developed with expert clinical input from primary care (2) and subspecialty (1) providers, as well as scientific input from individuals with training backgrounds in human factors engineering (1 PhD and 2 Masters). At the outset of each session, all participants underwent the same semistructured interview, including questions concerning experience and expertise with technology. With content tailored to their role (cancer survivor, caregiver, or provider), participants were then given use-case scenarios to perform using the PHR (Table 2) with a PC. The think-aloud design was used during scenario-testing and open-ended follow-up questions were used after each scenario. Concluding questions were then asked, with encouragement to envision a "blue-sky" future or the ideal PHR. The session was videotaped using Morae software so that verbal and visual cues and interactions could be analyzed.

Table 2. Use-case testing scenario, with an example of a colorectal cancer survivor.

Task	Description
1A	You recall that your doctor (Dr. Carter) wishes to know if you can grant him access to the electronic tool so he can view all of the information recorded there himself. He tells you to use the email: dcarter@fakeemail.com
1B	You decide you want to grant your spouse access to this electronic tool so she can view all of the information within it herself. Your spouse's email: myspouse@fakeemail.com
2A	Dr. Carter performed a colonoscopy on January 8, 2020, and found an abnormality, so your doctor has asked you to return in six months. Please record this information using the tool.
2B	After completing the last task, check to see when your next colonoscopy test will be and write your answer in the blank provided: Date of next colonoscopy test:_____
3A	After your doctor reviews your list of past treatments through this electronic tool, your doctor informs you of a mistake in the records in the electronic cancer toolkit. You were recorded as having been treated with Xeloda, but in reality you were treated with Erbitux. Use the system to update this piece of information.
3B	You want to share the radiation therapy you've received with your primary care doctor. Because your doctor cannot gain access to the toolkit, please write all the radiation treatment you've received in the spaces provided.
4	After updating your past treatment, please use the system to notify (via email) Greg Armstrong (Lance's Caregiver), whose email address is garmstrong@fakeemail.com
5	Use the system to record a personal experience involving what it's like to live with cancer. Please use this opportunity to express yourself, and please entitle it: "My Cancer Experience."

Analysis

For each role-based participant group, observations were analyzed using a grounded theory approach to determine the requirements for a CRC PHR and functions prioritized by each group. Two investigators, one of whom was not involved in the original interviews, performed the coding and analysis. The videotapes were broken into single, verbal or non-verbal, observations. The observations were recorded chronologically at the point in the session when they were collected and then grouped by participant role. Data analysis was inductive using the method of constant comparison, an iterative process consisting of an open and focused coding phase [22]. Within each participant role, the observations were coded using visual affinity diagramming [23] with an open-coding scheme. After independent coding by the investigators, coding discrepancies were resolved at consensus meetings. Investigators each read all transcripts and analyzed them for prevalent and recurrent themes. This phase elucidated three overarching themes: the web-based network, its functions, and its implementation. During the focused coding phase, investigators developed additional themes by conducting comparisons within and between transcripts, between themes, and finally across the role-based participant groups. Throughout the analysis, qualitative methods and procedures were used to ensure rigor and validity. These procedures included reflexivity (continually questioning interpretations and returning to the data to verify interpretations), search for alternative interpretations of the data, and depth of description (seeking out the specific details of participants' words) [24,25].

Results

Results Overview

All 6 patients were male, with an average age of 62.2 years (range 54-72 years); 2 of 6 (33%) were African American. Four caregiver spouses were recruited; all were women and half were African American. Of the health care providers, 4 of 7 (57%) were female, 5 (71%) were White, and 2 (29%) were Asian American. Overall, most patients would recommend the use of the PHR to other patients. Patient, caregiver, and provider observations could be grouped into three major themes: the network, its functions, and implementation. The "network" encompassed those who should be granted access to the PHR by the patient, "functions" comprised the helpful activities which the PHR enabled, and "implementation" included how best to implement the PHR into workflow and communication. These themes interrelate dynamically; for example, the types of individuals included in the network (provider, caregiver, or patient) may influence the range of possible uses.

During review, six additional themes emerged within these overarching themes: Network (Access Privileges and Communication), Functions (Self-tracking and Self-management; Journal (Reflection and Communication), and Implementation (Workflow and Future Enhancements).

Network

Access Privileges

The PHR allowed only the patient the ability to grant access to others; many permutations in who should be granted access were observed. All patients wanted physician access to their PHR. Most patients wanted family members to have access to the PHR; when family member access was granted, their spouse was always included. Access for family members was guided by relational closeness, and how important individual family members were to the patient's care. Without being prompted, 2 different patients suggested that researchers should also be provided access so as to create knowledge in the process of sharing information, expressing a sense of altruism: "probably use if I thought it would help other people."

Communication

Provider-to-provider communication: Participants from each stakeholder group saw the value of communication between providers. Patients and providers recommended that health information in the PHR be shared across health care systems. Patients also recommended sharing across different types of providers (both primary care and oncologist). One patient stated that primary care provider access should be "required."

Another patient suggested a doctor-only network, established without individual patient permissions. Patients suggested a model wherein medical personnel would have one-time access to PHR data in an emergency: "In the past, I had an uncle who died of diabetes and no one knew it but his wife, and she was out of town; he went to the emergency room and they gave him the wrong medicine."

Patient-to-provider communication: The PHR provided at least two opportunities for asynchronous communication between patients and providers through (1) secure email and (2) web-based journal access. Both patients and providers saw potential in the PHR for sharing medical information; patients more often saw potential in the tool for relationship-building. Each stakeholder group described how the optimal mode of communication varied by its purpose. Email was described as acceptable for simple messages, but not for complex or sensitive topics (eg, bad news) or issues requiring an immediate response, which were considered to be most appropriately communicated in person or by telephone.

Patient-to-caregiver communication: Among patients and caregivers, PHRs were perceived as adding value when family members lived at a distance from the patient. This distance could range from out-of-the-home to out-of-state. Patients discussed how providing caregivers access to the PHR could inform caregivers about possible side effects or the symptoms patients were experiencing. Although most spouses agreed with the desirability of access to the patient's PHR, one described their access to the patient's PHR as "an invasion of privacy."

Functions

Self-tracking and Self-management

Patients and caregivers saw the information recorded in the PHR as a useful memory tool because they "can't remember everything." One caregiver noted that this would be "great for keeping up with what is happening with 10 doctors." All stakeholder groups observed that the PHR would enable the patient and caregivers to record and track future testing, leading them to be "more engaged."

Patients and providers saw potential in the PHR for self-management. One patient said, "It would be nice to be able to keep up with how your treatment's going, kind of knowing if you're…getting better…A lot of times when you have cancer you always have a question mark over your head – where am I?" One provider discussed how the PHR could stimulate patients to ask questions. A caregiver noted that information from the PHR could be printed and brought to doctor appointments to address key issues. Patients and providers saw the potential for the PHR to track symptoms and perhaps deliver self-care approaches. Given the network of relationships that the PHR facilitated, patients saw an opportunity for collective self-management of their problems, facilitating help from caregivers.

Providers had reservations about patients viewing "raw data." They felt that patients were not prepared to interpret the data, which led to confusion and anxiety, and that inaccurate data in the record might upset patients. Patients and caregivers did not raise any downsides to having access to more information. Regarding inaccurate information in the PHR, patients expressed greater interest in how errors might be corrected rather than who would be to blame. None of the patients raised concerns about privacy or security of the PHR, but caregivers and providers expressed such concerns.

Journal: Reflection and Communication

Patients viewed the journal as a tool for reflection where they could record their personal thoughts, emotions, symptoms, and "vent" about frustrations. If shared, information recorded in the journal was seen as potentially reducing a sense of isolation: "Cancer is kind of a lonely illness to have; they can talk back and forth and share their experience of what they are going through and what medicines are working for them." Patients and caregivers saw the PHR as providing a way for others, including the health care provider, to better "understand the patient's issues".

In addition to self-reflection, patients and caregivers viewed the journal as a tool for communication and a way for patients to receive support from others. The "comments" section in the journal appeared to them to make bidirectional communication possible between patients and providers. For more stoic individuals, the journal was seen as offering a tool to communicate in written form what they might have trouble expressing verbally. One caregiver noted that a patient may withhold information about prognosis to "protect" family members, and the journal may enable greater sharing of information: "helps with hope if we know what to expect." Another caregiver suggested more multimedia resources in the PHR; for example, "songs or movies" that could help start conversations.

Two providers shared the view that the PHR would facilitate patient-physician communication, allow "sympathy," and help

the physician understand the patient "holistically." However, another provider worried that sharing journal access would make the doctor-patient relationship "less professional." Most providers were concerned about the time burden of processing a large amount of unstructured information. Potential malpractice liability owing to the provider having journal access was also raised, although a provider commented "you can't spend life worrying about lawsuits."

Implementation

Workflow

All stakeholder groups would prefer the PHR to be tethered to the patient's electronic health record (EHR), and did not see themselves as performing manual data entry. One participant indicated concern that only young or "techie" patients would be able to reliably use the PHR. For certain types of data, providers did not trust the accuracy or completeness of patient-entered information, although one provider stated that "patients should record their own values so they will be more involved."

The time burden of accessing the PHR was a common concern among providers. Strategies suggested for efficient adoption into workflow included nurse delegation; for example, email could be used for nurse-directed symptom management. Each stakeholder group believed the PHR should be well-integrated with other technologies to avoid creating multiple locations to access electronic health records or check email. PHR training was also perceived as necessary. Establishing parameters for patient uses and provider responses was considered good standard practice.

For email, providers were again concerned about the time burden but recognized that email could be both a "responsive" and "efficient" tool (eg, sending test results) for asynchronous communication. Patients expressed sensitivity to the time burdens of providers without prompting from the interviewer, suggesting that email may reduce the number of telephone calls, and expressing the opinion that email was more likely to reach their doctors.

Most providers considered email more efficient than the journal. A few providers indicated they would read the journal, but only if directed by patient request to a specific entry. For the journal, one provider suspected patients would record too much extraneous detail, thus making real issues harder to find. Providers suggested several types of structured information patients could enter, including review of systems, symptoms, and pain scores. Organizing tools, such as the use of subject headings and natural language processing, were suggested.

Future Enhancements

Patients and caregivers were interested in several specific enhancements to the PHR. They sought more guidance in accessing support groups and information regarding complementary medicine. Patients were interested not only in disease information, but also healthy lifestyle resources, especially nutrition. More capabilities concerning medication management were suggested, particularly a list of medications and side effects attributable to each chemotherapy agent. Other desired functions were the ability to refill medications, make appointments, and carry the PHR on portable devices.

Discussion

Principal Findings

The integration of multiple stakeholder perspectives regarding the potential use of a PHR for cancer survivors was a key strength of our study. Patients, caregivers, and providers all have unique roles and offer particular insights into the PHR's potential to meet the needs of cancer survivors. All stakeholder groups perceived the PHR to be a valuable tool and would recommend this patient-centered technology to others diagnosed with cancer. Several areas of agreement emerged across different stakeholder groups. First, the broader the network of users provided access, the better. As in other studies, a majority of patients wanted clinician access to the PHR [26], especially primary care physicians and oncologists. Stakeholder groups also recommended that networks bridge multiple health care systems. Essentially, participants were articulating a model of health information exchange, which shared electronic treatment information across multiple organizations, although perhaps they were unfamiliar with existing technical architectures or platforms to accomplish information-sharing [27].

Patient preference for access among individuals who were not clinicians was less universal and connected with the closeness of personal relationships and geographic proximity [8,28]. In patients with cancer, it may be especially beneficial for caregivers to be given access to the PHR [29], although such access needs to be balanced against the countervailing principles of patient privacy so as to prevent unwanted disclosures; for example, stigmatized conditions or billing information [30]. Overall, patients valued the ability to control access to the record on an individual basis [8]. While other systems allow the patient to control who else can access the PHR [31], this CRC PHR also enables the patient to control what domains of the records (medical versus personal) are accessed by whom. Consequently, the patient can share information with each individual provider and caregiver at the level that they choose. A patient could selectively provide access to the journal to family members owing to the personal nature of the content. Alternatively, a patient may choose to provide only medical providers with access to the treatment summary owing to its clinical nature and to preserve their privacy. But instead of making a priori assumptions about what decisions patient will make, the PHR provides patients with autonomy to tailor these decisions on the basis of their preferences for disclosure.

Another area of wide agreement among all stakeholder groups was the use of the PHR for information-sharing. Pragmatically, health information delivered through the PHR may increase patient recall and prompt questions at follow-up visits. By enabling patients to review their health information beforehand, and potential test recommendations, patients may be more prepared and activated [32,33] at physician visits. Furthermore, leveraging its information-sharing and communication functions, the PHR may serve as a foundation for collaborative decision-making and shared decisions.

Areas of disagreement were also noted among stakeholder groups, particularly between patients or caregivers and health care providers. Patient and caregivers both saw the value of the PHR in relationship-building. Information the patient shared about their personal cancer experience, especially through the journal function, was viewed as a way to be better understood as a whole person. Previous studies of narrative medicine suggest that patients' written stories of how illness has affected them can help them rediscover personal identity [34] and even improve patient outcomes [35]. Providers were concerned about the shift such web-based technologies could bring about in their professional roles. Previous research has outlined a mixed picture of social networking technologies. Social media use by patients led to more equal communication between the patient and provider, but increased doctor-switching [36]; patient-provider relationships may also be more harmonious owing to the opportunity to release negative emotions on the internet, but others have found suboptimal interactions between the patient and health care professional if providers do not agree with the information provided, or even feel their expertise challenged [37].

The messaging function, especially to communicate with their doctor, was highly valued by patients in this and other PHR studies [7,8,28,38]. Nonetheless, despite studies showing the feasibility of web-based patient-doctor communication [39], providers were concerned about time burden, data security, and privacy. While sensitive to time concerns, data privacy and security issues were not independently raised by the patients, which suggested that they were much more focused on the potential benefits than these mitigating risks. Consensus across stakeholder groups was easier to find on general guidelines for email use; for example, simple messages for nonemergent issues.

A related area of disagreement between patients and providers was the perceived value of sharing unstructured information. Patients saw the ability to construct meaning and relate personal experiences from sharing information in more qualitative forms, but only a minority of providers shared this view. Providers saw the same information as potentially inaccurate, a time burden, and source of medical liability. Structured information of all types was suggested by providers, including symptom and pain scores. Such discrete information may be more manageable for tracking and quality improvement purposes, as well as secondary research; however, the use of standardized instruments reduces the expressive content of the information and the patient's ability to articulate their unique circumstances.

Our findings deliver several key messages to be considered in the future design and implementation of patient-centered technologies.

PHRs Should Be Linked to EHRs

In free-standing PHRs, data entry would need to be performed by either the patient or the health care provider. Neither cancer survivors nor health care providers could see themselves as having adequate time to enter such data into the PHR. Moreover, health care providers did not have full confidence in the accuracy of patient-entered data. Based on the medical literature, their skepticism is warranted; for example, patients tend to overreport the receipt of cancer screening and underreport screening intervals [40]. Tethering or linking PHRs to EHRs would also enable the wide range of networking envisioned by patients. Providers who are already using EHRs could more readily be provided with access to PHRs. Through patient health information exchanges, multiple health care systems can be digitally connected. Broadband internet access serves as a barrier to PHR use among underserved populations and policy changes such as patient subsidies [41], and increased rural broadband infrastructure [42] also have a role to play in improving the health information ecosystem.

Emergency Care Is a Convincing Use Case

In this and other studies, patients expressed interest in allowing emergency care providers to temporarily access the PHR [43]. Cancer survivors and other chronic disease PHR users valued this emergency access, over privacy concerns [8,43]. A patient in this study and another study [8] related occasions where an unfavorable outcome resulted from the inability of emergency providers to access records. This patient-requested feature could be incorporated into future PHRs by a single-use key code, carried on the patient or held by an emergency contact.

Tailoring Is Essential

Tailoring across multiple dimensions of the PHR is possible, including access, content, and implementation approaches. As a matter of patient-centered principle, the patient was placed in control of PHR access; cancer survivors can then choose who to invite as well as what types of information those individuals can access. This flexible design appeared to be well-received, and patients reflected thoughtfully about to whom and why they would provide access. In our study and others, patients also wanted tailored guidance in searching for high-quality disease information and local support groups [44]. Of course, barriers to PHR adoption remain, and our group has previously identified barriers to use among a population of patients with CRC of similar age (mean age 58 years), including difficulty with system log-in, lack of computer literacy, and difficulty self-entering patient information [45].

The organizational contexts in which patients are seen, and in which PHRs will be administered, are quite heterogeneous. The US health care system has a multiplicity of practice environments, including academic and community, private and public, hospital- and office-based, and single and multispecialty groups. The structure and workflow of individual practices will play a large role in the optimal approach to implementation. Helpful guidance was suggested by our participants, including the use of support staff and best practices in the use of emails. Observations collected in multiple clinical settings would more fully inform other approaches worth consideration for dissemination.

Structured information was entered to summarize the patient's treatment and surveillance testing, and unstructured information was communicated in the journal and messaging system. Preferences for different types of information diverged between patients and providers. To resolve these varying perspectives, negotiation may be necessary between patients and providers to strike what both view as a fair balance. Different PHR

systems may be tailored over time to reflect these compromises in the types of information delivered and received.

Technology Changes Have the Potential to Change the Nature of the Patient-Physician Relationship

Patients almost universally valued the participation of health care providers in the PHR system. However, providers expressed reluctance about an open-ended engagement, especially in the case of patients expressing personal feelings or experiences in unstructured formats. While technology may initially appear to be a tool to facilitate patient-physician communication, it is worth considering how new tools such as the PHR could potentially change not only the mode, but also the content and qualities of communication. Broader sharing of the cancer experience may provide the opportunity for providers to more deeply understand their patients' identities. Further, the low adoption rate of PHRs [46] might be improved if these technologies enabled such high-quality, meaningful communication between patients and providers; a prior systematic review confirms that patients highly value using portals for communication with providers [47]. However, this more personal connection may narrow the professional distance between patients and providers in such a way that not all parties are comfortable. As more professional and social experiences transition to web-based digital platforms, the patient-physician relationship at the center of medicine may evolve in other unexpected or unintended ways. To our knowledge, our study results uniquely highlight the trade-offs and tensions that web-based technologies may introduce into the domain of patient-provider communication. Previous studies have perhaps focused upon the impact of shared records upon workflow [48], but not necessarily the nature of the patient-physician relationship itself. Our research design of incorporating both patient and provider perspectives was key to the discovery of these findings.

Limitations

This study was limited by its size, although owing to recurrent themes in the analysis, investigators believed that thematic saturation had been reached. Further, many of the main findings were consistent across different subject groups. The participant population was completely male; while males represent the strong majority of the US veteran population (>90%), female veterans should be aggressively recruited in future studies. The age distribution was also representative of patients diagnosed with colorectal cancer. Future studies should consider other patient groups and cancer types. Finally, no major EHR companies have deployed oncology patient portals that make possible the clinically tailored care delivered by the PHR tested here. Hence, we focused on general PHR issues (networking and implementation) and functions (self-management and journaling) relevant across portals. However, we believe that tailored, disease-focused PHRs have the potential to deliver greater clinical value to patients and providers; therefore, they represent a future model for technology design when the industry's business case can better support the degree of specialization required.

Conclusions

PHR perceptions are role-dependent, but there is marked consensus on many aspects of PHR design among stakeholders. This suggests that a single, integrated tool can be designed to meet several identified patient needs, including self-tracking and self-management, as well as more informed and shared medical decisions. Providers have unique concerns about the increased time burden and the accuracy of patient-entered data, and more fundamentally, how web-based communication tools may change the nature of the physician's professional role. Patients perceive these tools as a potential pathway to personal understanding that can deepen their relationships with doctors. Nonetheless, to realize this promise, patients and caregivers may need to search for and encourage health care providers to partner with them in exploring how emerging patient-centered technologies can be successfully implemented in modern medical practice to improve the relational quality of care.

Acknowledgments

The study was supported in part by VA HSR&D CDA 07-016, the VA Advanced Medical Informatics Fellowship Program (Carr), and the Livestrong Foundation. The funders had no role in the study design, collection, analysis and in the interpretation of data, writing of the report, or the decision to submit the paper for publication.

Conflicts of Interest

None declared.

References

1. Kim J, Jung H, Bates DW. History and trends of "Personal Health Record" research in PubMed. Healthc Inform Res 2011 Mar;17(1):3-17 [FREE Full text] [doi: 10.4258/hir.2011.17.1.3] [Medline: 21818452]
2. Cornbleet MA, Campbell P, Murray S, Stevenson M, Bond S, Joint Working Party of the Scottish Partnership Agency for PalliativeCancer CareNational Council for HospiceSpecialist Palliative Care Services. Patient-held records in cancer and palliative care: a randomized, prospective trialt. Palliat Med 2002 May;16(3):205-212. [doi: 10.1191/0269216302pm541oa] [Medline: 12046996]
3. Ko H, Turner T, Jones C, Hill C. Patient-held medical records for patients with chronic disease: a systematic review. Qual Saf Health Care 2010 Oct;19(5):e41. [doi: 10.1136/qshc.2009.037531] [Medline: 20511601]

4. Jones DA, Shipman JP, Plaut DA, Selden CR. Characteristics of personal health records: findings of the Medical Library Association/National Library of Medicine Joint Electronic Personal Health Record Task Force. J Med Libr Assoc 2010 Jul;98(3):243-249 [FREE Full text] [doi: 10.3163/1536-5050.98.3.013] [Medline: 20648259]
5. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc 2011;18(4):515-522 [FREE Full text] [doi: 10.1136/amiajnl-2011-000105] [Medline: 21672914]
6. Consumers and Health Information Technology: A National Survey. California Health Care Foundation. 2010 Apr 13. URL: http://www.chcf.org/publications/2010/04/consumers-and-health-information-technology-a-national-survey [accessed 2021-12-27]
7. De Clercq PA, Hasman A, Wolffenbuttel BHR. A consumer health record for supporting the patient-centered management of chronic diseases. Med Inform Internet Med 2003 Jun;28(2):117-127. [doi: 10.1080/14639230310001600506] [Medline: 14692588]
8. Lafky D, Horan T. Prospective Personal Health Record Use Among Different User Groups: Results of a Multi-wave Study. 2008 Presented at: 41st Annual Hawaii International Conference on System Sciences (HICSS 2008); January 7-10, 2008; Waikoloa, HI. [doi: 10.1109/hicss.2008.363]
9. Sox CM, Gribbons WM, Loring BA, Mandl KD, Batista R, Porter SC. Patient-centered design of an information management module for a personally controlled health record. J Med Internet Res 2010 Aug 30;12(3):e36 [FREE Full text] [doi: 10.2196/jmir.1269] [Medline: 20805091]
10. Atreja A, Mehta N, Miller D, Moore S, Nichols K, Miller H, et al. One size does not fit all: using qualitative methods to inform the development of an Internet portal for multiple sclerosis patients. AMIA Annu Symp Proc 2005:16-20 [FREE Full text] [Medline: 16778993]
11. Jemal A, Simard EP, Dorell C, Noone A, Markowitz LE, Kohler B, et al. Annual Report to the Nation on the Status of Cancer, 1975-2009, featuring the burden and trends in human papillomavirus(HPV)-associated cancers and HPV vaccination coverage levels. J Natl Cancer Inst 2013 Feb 06;105(3):175-201 [FREE Full text] [doi: 10.1093/jnci/djs491] [Medline: 23297039]
12. National Research Council, Institute of Medicine. In: Ganz PA, Hewitt PM, editors. From Cancer Patient to Cancer Survivor: Lost in Transition: An American Society of Clinical Oncology and Institute of Medicine Symposium. Washington, DC: National Academies Press; 2006.
13. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. CA Cancer J Clin 2020 May;70(3):145-164 [FREE Full text] [doi: 10.3322/caac.21601] [Medline: 32133645]
14. Rex DK, Kahi CJ, Levin B, Smith RA, Bond JH, Brooks D, American Cancer Society, US Multi-Society Task Force on Colorectal Cancer. Guidelines for colonoscopy surveillance after cancer resection: a consensus update by the American Cancer Society and the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 2006 May;130(6):1865-1871. [doi: 10.1053/j.gastro.2006.03.013] [Medline: 16697749]
15. Ansari N, Solomon MJ, Fisher RJ, Mackay J, Burmeister B, Ackland S, et al. Acute adverse events and postoperative complications in a randomized trial of preoperative short-course radiotherapy versus long-course chemoradiotherapy for T3 adenocarcinoma of the rectum: Trans-Tasman Radiation Oncology Group Trial (TROG 01.04). Ann Surg 2017 May;265(5):882-888. [doi: 10.1097/SLA.0000000000001987] [Medline: 27631775]
16. Gebremedhn EG, Shortland PJ, Mahns DA. The incidence of acute oxaliplatin-induced neuropathy and its impact on treatment in the first cycle: a systematic review. BMC Cancer 2018 Apr 12;18(1):410 [FREE Full text] [doi: 10.1186/s12885-018-4185-0] [Medline: 29649985]
17. Nazi KM. The personal health record paradox: health care professionals' perspectives and the information ecology of personal health record systems in organizational and clinical settings. J Med Internet Res 2013 Apr 04;15(4):e70 [FREE Full text] [doi: 10.2196/jmir.2443] [Medline: 23557596]
18. Tieu L, Sarkar U, Schillinger D, Ralston JD, Ratanawongsa N, Pasick R, et al. Barriers and facilitators to online portal use among patients and caregivers in a safety net health care system: a qualitative study. J Med Internet Res 2015 Dec 03;17(12):e275 [FREE Full text] [doi: 10.2196/jmir.4847] [Medline: 26681155]
19. Casillas A, Cemballi AG, Abhat A, Lemberg M, Portz JD, Sadasivaiah S, et al. An untapped potential in primary care: semi-structured interviews with clinicians on how patient portals will work for caregivers in the safety net. J Med Internet Res 2020 Jul 20;22(7):e18466 [FREE Full text] [doi: 10.2196/18466] [Medline: 32706709]
20. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. 1993 Presented at: ACM INTERCHI'93 Conference; April 24-29, 1993; Amsterdam, The Netherlands p. 206-213 URL: https://dl.acm.org/doi/10.1145/169059.169166
21. Seebregts CJ, Mamlin BW, Biondich PG, Fraser HSF, Wolfe BA, Jazayeri D, OpenMRS Implementers Network. The OpenMRS Implementers Network. Int J Med Inform 2009 Nov;78(11):711-720. [doi: 10.1016/j.ijmedinf.2008.09.005] [Medline: 19157968]
22. Charmaz K. Constructing Grounded Theory: A practical guide through qualitative analysis. Thousand Oaks, CA: Sage; 2006.
23. Brassard M. Featuring the Seven Management and Planning Tools. In: The Memory Jogger Plus+. Massachusetts: Goal/Qpc; 1989.

24. Borkan J. Immersion/Crystallization. In: Crabtree BF, Miller WL, editors. Doing Qualitative Research (2nd edition). Thousand Oaks, CA: Sage Publications; 1999:179-194.
25. Davies D, Dodd J. Qualitative research and the question of rigor. Qual Health Res 2002 Feb;12(2):279-289. [doi: 10.1177/104973230201200211] [Medline: 11837376]
26. Schwartz PH, Caine K, Alpert SA, Meslin EM, Carroll AE, Tierney WM. Patient preferences in controlling access to their electronic health records: a prospective cohort study in primary care. J Gen Inter Med 2015 Jan;Suppl 1(Suppl 1):25-30. [doi: 10.1007/s11606-014-3054-z] [Medline: 25480721]
27. Medford-Davis LN, Chang L, Rhodes KV. Health Information Exchange: What do patients want? Health Informatics J 2017 Dec;23(4):268-278 [FREE Full text] [doi: 10.1177/1460458216647190] [Medline: 27245671]
28. Zulman DM, Nazi KM, Turvey CL, Wagner TH, Woods SS, An LC. Patient interest in sharing personal health record information: a web-based survey. Ann Intern Med 2011 Dec 20;155(12):805-810 [FREE Full text] [doi: 10.7326/0003-4819-155-12-201112200-00002] [Medline: 22184687]
29. Hogarth M, Hajopoulos K, Young M, Cowles N, Churin J, Hornthal B, et al. The Communication and Care Plan: a novel approach to patient-centered clinical information systems. J Biomed Inform 2010 Oct;43(5 Suppl):S6-S8 [FREE Full text]
30. Latulipe C, Quandt SA, Melius KA, Bertoni A, Miller DP, Smith D, et al. Insights into older qdult patient concerns qround the caregiver proxy portal use: qualitative interview study. J Med Internet Res 2018 Nov 02;20(11):e10524 [FREE Full text] [doi: 10.2196/10524] [Medline: 30389654]
31. Leventhal JC, Cummins JA, Schwartz PH, Martin DK, Tierney WM. Designing a system for patients controlling providers' access to their electronic health records: organizational and technical challenges. J Gen Intern Med 2015 Jan;30 Suppl 1:S17-S24 [FREE Full text] [doi: 10.1007/s11606-014-3055-y] [Medline: 25480722]
32. van Dulmen S, Driesenaar JA, van Weert JC, van Osch M, Noordman J. PatientVOICE: development of a preparatory, pre-chemotherapy online communication tool for older patients with cancer. JMIR Res Protoc 2017 May 10;6(5):e85 [FREE Full text] [doi: 10.2196/resprot.6979] [Medline: 28490421]
33. Epping-Jordan JE, Pruitt SD, Bengoa R, Wagner EH. Improving the quality of health care for chronic conditions. Qual Saf Health Care 2004 Aug;13(4):299-305 [FREE Full text] [doi: 10.1136/qhc.13.4.299] [Medline: 15289634]
34. Slocum RB, Howard TA, Villano JL. Narrative Medicine perspectives on patient identity and integrative care in neuro-oncology. J Neurooncol 2017 Sep;134(2):417-421. [doi: 10.1007/s11060-017-2542-5] [Medline: 28669013]
35. Cepeda MS, Chapman CR, Miranda N, Sanchez R, Rodriguez CH, Restrepo AE. Emotional disclosure through patient narrative may improve pain and well-being: results of a randomized controlled trial in patients with cancer pain. J Pain Symptom Manage 2008 Jun;35(6):623-631. [doi: 10.1016/j.jpainsymman.2007.08.011] [Medline: 18359604]
36. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: A systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug;26(1):442. [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]
37. Rupert DJ, Moultrie RR, Read JG, Amoozegar JB, Bornkessel AS, O'Donoghue AC. Perceived healthcare provider reactions to patient and caregiver use of online health communities. Patient Educ Couns 2014 Sep;96(3):320-326. [doi: 10.1016/j.pec.2014.05.015] [Medline: 24923652]
38. Segall N, Saville JG, L'Engle P, Carlson B, Wright MC, Schulman K, et al. Usability evaluation of a personal health record. AMIA Annu Symp Proc 2011;2011:1233-1242 [FREE Full text] [Medline: 22195184]
39. Ye J, Rust G, Fry-Johnson Y, Strothers H. E-mail in patient-provider communication: a systematic review. Patient Educ Couns 2010 Aug;80(2):266-273 [FREE Full text] [doi: 10.1016/j.pec.2009.09.038] [Medline: 19914022]
40. Gordon NP, Hiatt RA, Lampert DI. Concordance of self-reported data and medical record audit for six cancer screening procedures. J Natl Cancer Inst 1993 Apr 07;85(7):566-570. [doi: doi: 10.1093/jnci/85.7.566] [Medline: 8455203]
41. Perzynski AT, Roach MJ, Shick S, Callahan B, Gunzler D, Cebul R, et al. Patient portals and broadband internet inequality. J Am Med Inform Assoc 2017 Sep 01;24(5):927-932 [FREE Full text] [doi: 10.1093/jamia/ocx020] [Medline: 28371853]
42. Cortelyou-Ward K, Atkins DN, Noblin A, Rotarius T, White P, Carey C. Navigating the digital divide: barriers to telehealth in rural areas. J Health Care Poor Underserved 2020;31(4):1546-1556. [doi: 10.1353/hpu.2020.0116] [Medline: 33416736]
43. Walker J, Ahern DK, Le LX, Delbanco T. Insights for internists: "I want the computer to know who I am". J Gen Intern Med 2009 Jun;24(6):727-732 [FREE Full text] [doi: 10.1007/s11606-009-0973-1] [Medline: 19412641]
44. Tenforde M, Jain A, Hickner J. The value of personal health records for chronic disease management: what do we know? Fam Med 2011 May;43(5):351-354 [FREE Full text] [Medline: 21557106]
45. Tarver WL, Robb BW, Haggstrom DA. Usefulness and usability of a personal health record and survivorship care plan for colorectal cancer survivors: survey study. JMIR Cancer 2019 Aug 20;5(2):e10692 [FREE Full text] [doi: 10.2196/10692] [Medline: 31432780]
46. Fraccaro P, Vigo M, Balatsoukas P, Buchan IE, Peek N, van der Veer SN. Patient portal adoption rates: a systematic literature review and meta-analysis. Stud Health Technol Inform 2017;245:79-83. [Medline: 29295056]
47. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015 Feb 10;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
48. Cajander Å, Moll J, Englund S, Hansman A. Medical records online for patients and effects on the work environment of nurses. Stud Health Technol Inform 2018;247:271-275. [Medline: 29677965]

Abbreviations

CRC: colorectal cancer
EHR: electronic health record
IOM: Institute of Medicine
PHR: personal health record
RVAMC: Roudebush Veterans Affairs Medical Center

Edited by A Kushniruk; submitted 30.09.19; peer-reviewed by J Alpert, T Schopf, J Wu; comments to author 09.06.20; revised version received 07.02.21; accepted 11.09.21; published 25.01.22.

Please cite as:
Haggstrom DA, Carr T
Uses of Personal Health Records for Communication Among Colorectal Cancer Survivors, Caregivers, and Providers: Interview and Observational Study in a Human-Computer Interaction Laboratory
JMIR Hum Factors 2022;9(1):e16447
URL: https://humanfactors.jmir.org/2022/1/e16447
doi:10.2196/16447
PMID:35076406

Original Paper

An Information Directory App InHouse Call for Streamlining Communication to Optimize Efficiency and Patient Care in a Hospital: Pilot Mixed Methods Design and Utility Study

George Schilling[1], DO; Leonardo Villarosa[2], MD

[1]Vidant Medical Center, Greenville, NC, United States

[2]Vidant Medical Center/East Carolina University, Greenville, NC, United States

Corresponding Author:
George Schilling, DO
Vidant Medical Center
2100 Statonsburg Rd
Greenville, NC, 27834
United States
Phone: 1 252 847 4441
Fax: 1 252 847 6086
Email: georgeschilling@gmail.com

Abstract

Background: Communication failures disrupt physician workflow, lead to poor patient outcomes, and are associated with significant economic burden. To increase efficiency when contacting a team member in a hospital, we have designed an information directory app, InHouse Call.

Objective: This study aimed to describe the design of InHouse Call, objectively compare the usefulness of the app versus that of traditional methods (operator or pocket cards, etc), and determine its subjective usefulness through user surveys and a net promoter score (NPS).

Methods: This pilot study utilizing before-after trials was carried out at a tertiary academic hospital and involved 20 clinicians, including physiatrists, hospitalists, internal medicine and family medicine residents, and advanced practice providers/nurse practitioners/physician assistants. InHouse Call was designed to efficiently supply contact information to providers through a simple, user-friendly interface. The participants used InHouse Call in timed trials to contact a health care team member in the hospital via a telephone call. The effectiveness of InHouse Call in connecting the user with a contact in the hospital was measured through timed trials comparing the amount of time spent in attempting to make the connection using traditional methods versus the app. Usability was measured through exit surveys and NPS.

Results: The average time spent connecting to the correct contact using traditional methods was 59.5 seconds, compared to 13.8 seconds when using InHouse Call. The degree of variance when using traditional methods was 1544.2, compared to 19.7 with InHouse Call. A call made using the traditional methods deviated from the mean by 39.3 seconds, compared to 4.4 seconds when using InHouse Call. InHouse Call achieved an NPS of 95.

Conclusions: InHouse Call significantly reduced the average amount of time spent connecting with the correct contact as well as the variability to complete the task, thus proving to be the superior method of communication for health care providers. The app garnered a high NPS and positive subjective feedback.

(JMIR Hum Factors 2022;9(1):e23833) doi:10.2196/23833

KEYWORDS

InHouse Call; communication; hospital directory; healthcare; health care; health informatics; mHealth; mobile app; digital health; patient records; electronic health

Introduction

Background

The current fast-paced field of health care requires frequent communication among all health care team members including providers, case managers, nursing staff, therapists, pharmacists, nutritionists, technicians, and more. Owing to rapid expansion, hospitals in the United States often have their communications systems partitioned as opposed to a unified structure. Thus, communication failures have been a main source of concern in poor patient outcomes and often identified as a root cause of fatal errors [1]. In fact, entire simulation-based trainings have been dedicated to improving physician communication [2]. Although there have been several technological advancements in hospital communication in the past several decades, they have often acted as mere add-ons to an already complex communication system [3], and there has been limited evidence for improvement in effective interprofessional communication [4]. The time lost owing to poor hospital communication resulting in delayed testing and increased admission days comes at a significant economic burden with yearly estimates ranging US $12-$30 billion [5,6].

In large hospitals, there are two traditional methods used to contact a team member: (1) calling the operator and waiting to be connected and (2) pocket cards or workstation sheets with lists of general numbers. Many health care providers associate the operator method with long wait times and dropped or incorrectly transferred calls. While small and compact, pocket cards take up space in an already cluttered coat pocket, and their type font is often small and difficult to read. Both pocket cards and workstation sheets are physically limited in terms of the amount of information they can contain and sometimes have outdated information. Lamination of pocket cards prevents addition of information, while scribbling additional numbers onto the margins of workstation sheets leads to confusion.

Time-motion studies since the late 1980s have shown the downward trend of time spent in direct patient care and more time spent talking with physicians [7,8]. More recent studies have shown a sharper decline with the advent of the electronic health record system [9]. As health care becomes more complex, it is imperative to optimize and streamline time-consuming processes.

Technology is changing the face of health care. Physicians who adapt and work together with technology find success in a rapidly evolving system, and how well a physician adapts to a new technology is valuable to its success [10].

Goals of This Intervention

To increase provider efficiency with contacting a team member in a large hospital, we designed and implemented a novel native information directory prototype app, InHouse Call. InHouse Call is a simple hospital directory tool used to facilitate seamless communication between providers and among health care team members by eliminating call transfers and making available the most up-to-date contact information for users. The user interface and back end of the app prioritizes ease of use and speed for completing the task.

Research Goals

In this paper, we describe the design and future development of InHouse Call, objectively compare the usefulness of the app versus traditional methods via timed trials, and determine its subjective usefulness through user surveys and its net promoter score (NPS). The NPS is a standardized scoring system that rates the likelihood of a user recommending a new product or technology to a colleague [11].

Methods

Study Design and Setting

In this mixed methods study, we evaluated InHouse Call using a before-and-after survey methodology combined with objective timed trials. This study was performed at an academic medical center with a mix of 20 internal medicine and family medicine residents, hospitalists, and other inpatient clinicians.

Intervention: InHouse Call

InHouse Call was developed by the author, a physical medicine and rehabilitation resident, using an iterative, user-centric interface with a focus on usability and efficiency [12]. The creation of the app was inspired by the difficulty experienced attempting to contact the hospital's echocardiogram (ECHO) department while trying to complete a syncope work-up. This event resulted in an unnecessarily prolonged stay for a patient. Similar experiences shared by colleagues ultimately led to the development of InHouse Call.

InHouse Call is a native mobile-based software app designed to supply the correct contact information to providers in the most efficient manner possible with a simple user-friendly interface (Figure 1).

InHouse Call is designed to contain the comprehensive database of health care team contact information found in large hospitals, including nursing staff, charge nurses, unit secretaries, case managers, pharmacists, nutritionists, and others, all searchable by the patient's room number. The app also contains important department numbers, such as the radiology and laboratory departments, as well as hospital administration and clinic contact information. InHouse Call integrates seamlessly with the preexisting telephone system, as opposed to alternative health care messaging apps that work in a closed loop. It is important to note that InHouse Call is completely Health Insurance Portability and Accountability Act (HIPAA)–compliant as no personally identifiable or protected health information of any party is stored either locally or remotely.

In our app design, the user opens the app leading directly to a home screen providing the main "Search Patient Room" searchable database as well as six subfolders containing pertinent, easy-to-read information, including the following: Units, Departments, Clinics, Admin Numbers, Misc, and About. Unique to the InHouse Call design, the "Search Patient Room" database allows the user to enter the patient room number without requiring patient-identifiable information, which then provides fixed health care team member information assigned to that specific patient room. This includes the registered nurse (RN) pod or hallway phone, case manager, or pharmacist who

is directly assigned to that room. It is important to note that the team member's personal identifiable information is not included; only the phone number assigned to that team member is provided, thus ensuring HIPAA compliance.

The app is currently in the native form and created with Android Studio and written in Java. Only textual data—no images or video—are stored on-premise in a SQLite database built into the app, using a dual-encrypted secure server. This design was chosen to ensure the highest-speed search output available in large hospital center environments where internet connectivity may not be 100% reliable, especially in corridors or elevators where physicians are often located while travelling between patient rooms. Although the current system is in a native app, Figure 2 describes the web app phase where the data will be stored on a server but still cached on the native app to ensure the same level of speed.

In the anticipated archistructure described in Figure 2, there are several noteworthy integration points for the future planned software. First, the user login and authentication process serve two purposes: (1) it provides security for the hospital contact information available, and (2) it allows a single provider to switch between different hospital directories. Clinicians often work at different hospitals, sometimes even between different health care systems, and this feature will allow them to log out of one hospital directory and log in to another hospital directory. Next, the download or cache step ensures that the most up-to-date information is available in the directory. This step is designed to balance speed with functionality. A key anticipated feature also includes calls being directly made from the user's smartphone containing the app.

Figure 1. InHouse Call homepage with a searchable database and contact folders.

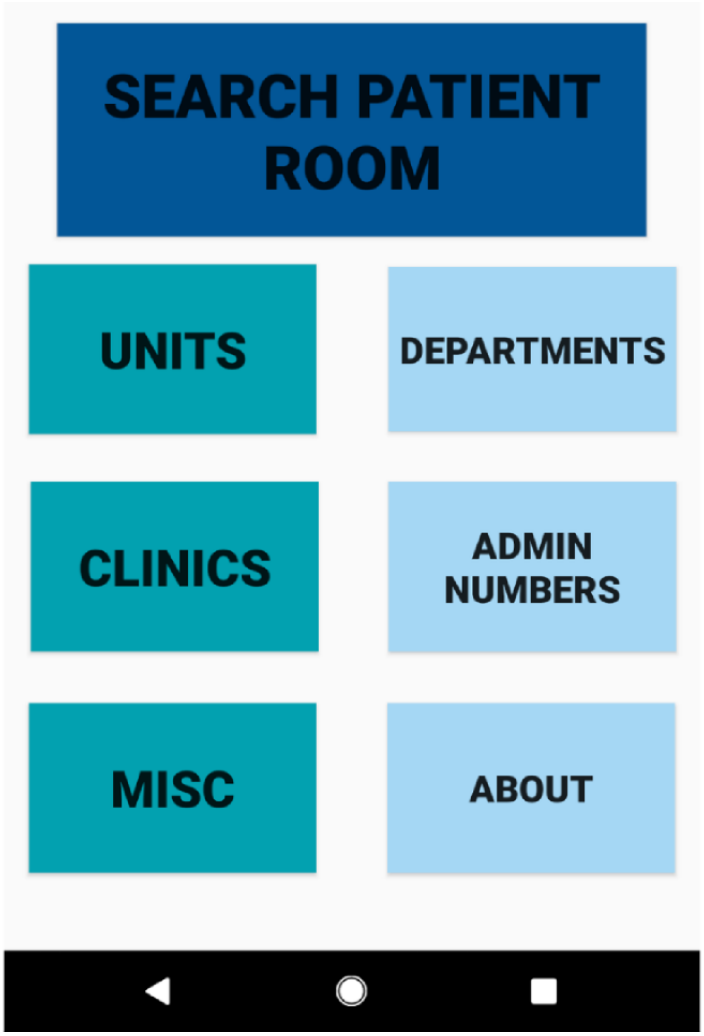
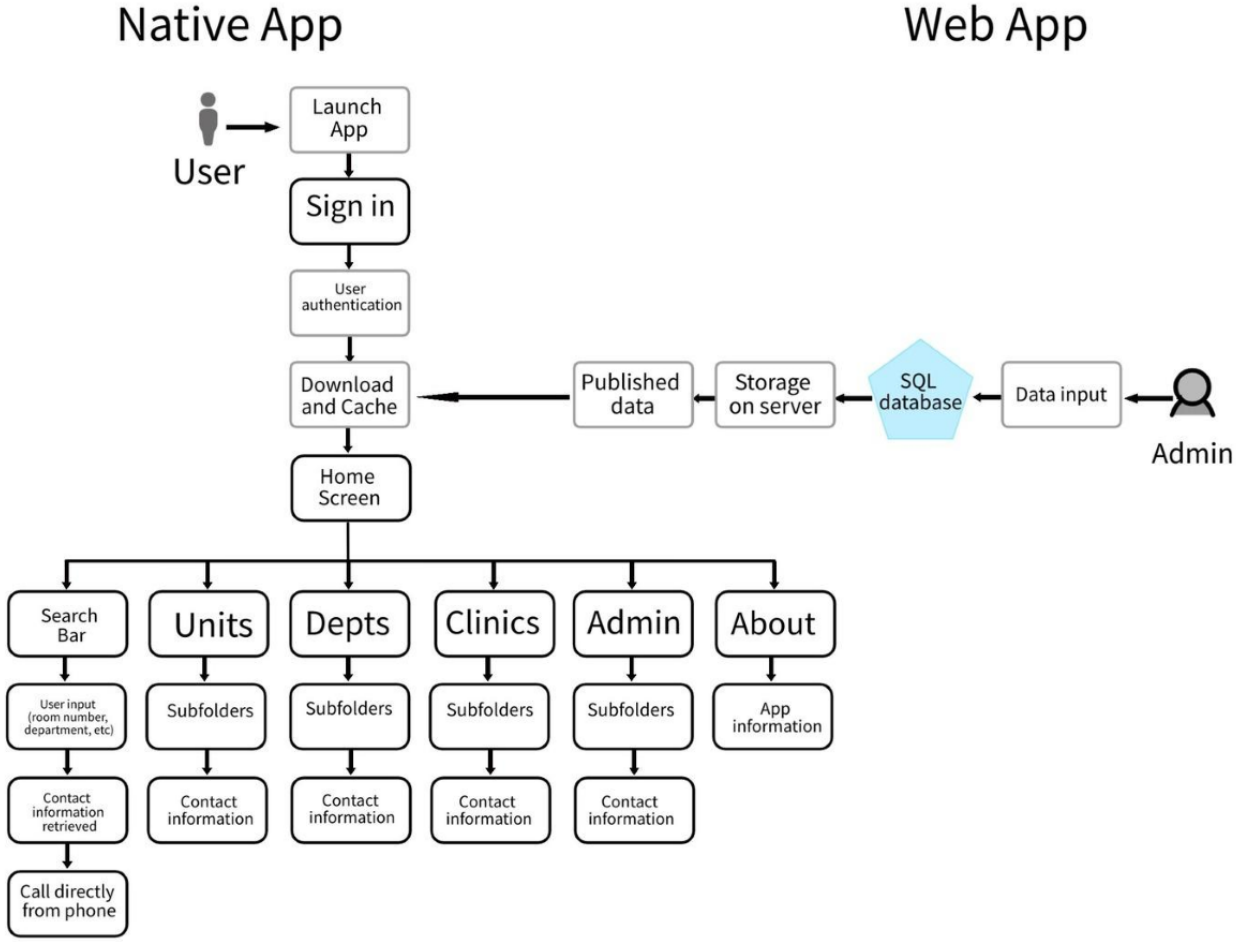


Figure 2. InHouse Call sitemap. Anticipated integration of a web app data input system to keep InHouse Call updated.



Participant Selection

In total, 20 providers were surveyed and participated in the timed trials. The participants mainly included hospitalists and internal medicine and family medicine residents as they were found to require the most communication with health care team members. The study also included a large variety of clinicians who mainly work in the hospital setting, including surgery residents and nurse practitioners (NPs) or physician assistants (PAs), intensivists, pediatricians, endocrinologists, physiatrists, and emergency medicine and obstetrics and gynecology residents. This was done to test the usefulness of the app by comparing providers who had high and low call volumes.

Study Protocol

Before the introduction of InHouse Call, the participants completed a baseline survey of their experience with the current hospital communication system, including its impact on their workflow and how much time they spent searching for the correct contacts each day. All participants completed the baseline survey. After the completion of the survey, the clinicians then participated in 2 rounds of timed trials, one using their preferred traditional methods of either calling the operator or using a pocket card/workstation contact list, and one using InHouse Call.

For the first round of the timed trial, a piece of paper was placed face down on the table in front of the participant at their workstation, and the participant was instructed to complete the task outlined on the piece of paper; namely, to make a phone call either to a case manager, nurse, department, etc. The participant was told that the timer would start when the piece of paper was turned over and end when the call to the correct contact rang once. The contacts were made aware ahead of time that their phones may be ringing temporarily.

Before the second round of the timed trial, the participants were then briefly introduced to InHouse Call. They were given 45-60 seconds to become familiar with its functionality. They were shown how to locate a contact number using the “Search By Patient Room” feature as well as by the department and clinic subfolders. The same trials were conducted again with a smartphone placed on the workstation table with the screen off and app closed. These providers were instructed to again follow the prompts on the piece of paper in front of them, this time using InHouse Call to complete the task. In total, 85% of the participants completed the trial phase of the study. The providers who declined to participate in the timed trials were still introduced to InHouse Call and its functionality.

The final phase of the study was an exit survey. Although only 85% of the participants completed the trial phase of the study, 100% of participants were shown and interacted with InHouse Call and completed the exit survey, which included the NPS questionnaire. Survey respondents were also asked about their comfort level with the app to determine its usability.

Outcome Measures

We evaluated InHouse Call's effectiveness in connecting the user with the correct contact in the hospital through the use of objective timed trials and comparing the amount of time spent in attempting to make the connection using traditional methods (operator or pocket cards, etc) versus InHouse Call. Usability was further quantified through the use of the NPS and the exit survey.

Analysis Approach

We classified the participants by their area of practice, their number of calls made in a workday, and their experience with the current communication system. We then compared the results of the timed trials using the traditional communication methods versus InHouse Call. The results of the two methods were further compared by average time, SD, and variance. The NPS was obtained to determine the successful rollout of the app, and the feedback obtained in the exit survey was categorized into general themes or categories.

Qualitative data were managed and analyzed using Google Sheets (Alphabet Inc).

Results

Characteristics of Study Subjects

There was a 100% response rate in the entrance and exit survey for the 20 participants, while 85% of participants completed some or all of the timed trials. Participants were classified in accordance with the area of practice and number of calls made in an average work day (Tables 1 and 2).

Table 1. Composition of the participating health care providers.

Provider specialty	Proportion, %
Obstetrics and gynecology	5
Emergency medicine	5
Pediatrics	5
Neonatal intensive care unit nurse practitioner	5
Endocrinology	5
Physical medicine and rehabilitation	10
Surgery	20
Internal medicine, family medicine, or hospitalist	45

Table 2. Average number of calls made per participant.

Calls, n	Participants, n (%)
0-5	1 (5)
6-10	7 (35)
11-15	5 (25)
16-20	4 (20)
21-25	1 (5)
26-30	2 (10)

Survey Results

In the entrance survey, 55% of respondents reported that it took over a minute to connect with a health care team member through the operator, while 65% reported the frequency of dropped calls or being transferred to the wrong person when using the operator as either occasionally or frequently (Tables 3 and 4).

In total, 75% of respondents reported frustration over not being able to find the right contact daily or several times per week, and 70% reported that poor communication impacted patient care and workflow daily or several times per week (Table 5). None responded that they were never frustrated by the inability to contact the right person, or that poor communication never impacted patient care and workflow.

A total of 80% of respondents reported spending at least 5 minutes of their workday searching for the right contact, with 10% reporting spending at least 20 minutes of their workday (Table 6). When asked to determine the number 1 complaint with communication in their hospital, 55% of respondents reported "Hard to find the right number" (Table 7).

Table 3. Perceived time spent with the operator to determine how long respondents felt it would take them to connect with a health care team member (registered nurse, case manager, etc) through the switchboard.

Self-reported perceived time (minutes)	Participants (n=16)[a], n (%)
<0.5	0 (0)
0.5-1	3 (15)
<1	2 (10)
1-1.5	6 (30)
1.5-2	3 (15)
2.2-5	0 (0)
>2.5	2 (10)

[a]Drop-out rate=20% (n=4 participants).

Table 4. Frequency of wrong transfers or dropped calls among respondents when using the switchboard method to reach a health care provider.

Frequency	Respondents (n=18)[a], n (%)
Never	0 (0)
Very rarely	2 (10)
Rarely	3 (15)
Occasionally	6 (30)
Frequently	7 (35)
Always	0 (0)

[a]Drop-out rate=10% (n=2 participants).

Table 5. Frequency of frustration among respondents (N=20) on not being able to find the right contact and poor communication affecting patient care and workflow.

Frequency	Respondents frustrated on not finding the right contact, n (%)	Poor communication affecting care delivery and workflow among respondents, n (%)
Never	0 (0)	0 (0)
Once per month	0 (0)	1 (5)
Several times per month	2 (10)	4 (20)
Once per week	3 (15)	1 (5)
Several times per week	8 (40)	8 (40)
Daily	7 (35)	6 (30)

Table 6. Time spent by respondents (N=20) in searching the right contact on each day.

Time (minutes)	Respondents, n (%)
0-5	4 (20)
5-10	7 (35)
10-20	7 (35)
20-30	1 (5)
30-60	1 (5)
>60	0 (0)

Table 7. Primary complaint of respondents (N=20) with communication at the hospital.

Complaint	Respondents, n (%)
Transferred to the wrong person	1 (5)
Takes too much time	5 (25)
Difficult to find the right number	11 (55)
All three circled	1 (5)
Unable to reach intended person	1 (5)
Poor interdepartmental communication	1 (5)

Timed Trials Results

Of the 20 survey respondents, 75% participated in the timed trial to reach an RN. The average participant spent 78.7 seconds to reach an RN via traditional methods and spent only 16 seconds when using InHouse Call (Table 8). Of the 15 trials using traditional methods, one participant (participant 8) gave up after 2 minutes owing to incomplete information available on their pocket card, while another (participant 12) spent 203 seconds connecting to the right contact after experiencing a dropped call.

Of the 20 survey respondents, 60% participated in the timed trial to reach the ECHO department. The average participant spent 49.9 seconds to reach the ECHO department via traditional methods and spent 11.1 seconds via InHouse Call (Table 8). Of the 12 trials using traditional methods, 3 participants gave up (participants 6, 10, and 11): one owing to incomplete information available on their pocket card, one owing to long wait times, and the last one owing to a wrong transfer.

Of the 20 survey respondents, 11 (55%) participated in the timed trial to reach a clinic. The average participant spent 43.9 seconds to reach a clinic via traditional methods and spent 13.8 seconds via InHouse Call (Table 8). Of the 11 trials using traditional methods, 2 participants (participants 6 and 10) gave up owing to incomplete information available on their pocket card.

Of the 20 survey respondents, 6 (30%) participated in the timed trial to reach a wound care team member. All participants were unable to complete the task using traditional methods, averaging 161.8 seconds before giving up and averaging 19.2 seconds via InHouse Call with success (Table 8). The longest time a participant attempted to reach a wound care team member using traditional methods was a little over 300 seconds.

Out of 44 trials in total, the average time spent to connect to the correct contact, via traditional methods was 73.5 seconds, while the average time spent to connect to the correct contact via InHouse Call was 14.6 seconds (Table 9). As no participant was able to connect with a wound care team member via traditional methods, that set of data collected may be considered outliers. With the wound care team outlier data removed, the average time spent to connect to the correct contact via traditional methods was 59.5 seconds while the average time spent via InHouse Call was 13.8 seconds.

Also shown in Table 9 is the degree of variability between each call using traditional methods versus the consistency achieved using InHouse Call. Analyzing the data without outliers, the degree of variance using traditional methods was 1544.2 compared to 19.7 with InHouse Call. A call made using the traditional methods deviated from the mean by 39.3 seconds, while a call using InHouse Call deviated from the mean by 4.4 seconds.

Table 8. Time to reach different health care providers by using the traditional method versus InHouse Call.

Participant	Time to reach registered nurses (seconds)[a]		Time to reach the echocardiogram department (seconds)[b]		Time to reach a clinic (seconds)[c]		Time to reach the wound care team (seconds)[d]	
	Traditional method	InHouse Call	Traditional method	InHouse Call	Traditional method	InHouse Call	Traditional method	InHouse Call
1	100	20	18	10	24	15	62	20
2	48	11	70	14	58	25	121	30
3	55	17	52	10	62	15	123	12
4	33	20	50	19	45	15	180	22
5	37	17	15	5	30	10	183	16
6	25	20	63	12	61	12	302	15
7	85	10	100	10	51	14	N/A[e]	N/A
8	127	14	20	12	15	15	N/A	N/A
9	68	20	41	10	25	10	N/A	N/A
10	47	12	120	13	60	13	N/A	N/A
11	58	12	17	11	52	8	N/A	N/A
12	203	16	33	7	N/A	N/A	N/A	N/A
13	124	20	N/A	N/A	N/A	N/A	N/A	N/A
14	43	10	N/A	N/A	N/A	N/A	N/A	N/A
15	128	22	N/A	N/A	N/A	N/A	N/A	N/A

[a]Two participants gave up on the task on using traditional methods.

[b]Three participants gave up on the task on using traditional methods.

[c]Two participants gave up on the task on using traditional methods.

[d]All 6 participants gave up on the task on using traditional methods.

[e]N/A: not applicable.

Table 9. Comparison of timed trials by average time, SD, and variance.

	Average time	Mean deviation	Variance (SD)	*t* test (*df*)
Traditional methods	73.50	42.75	3328.52 (57.69)	<0.001 (43)
InHouse Call	14.57	3.98	25.02 (5.00)	N/A[a]
Traditional methods without wound team data	59.55	28.28	1544.19 (39.30)	<0.001 (4)
InHouse Call without wound team data	13.84	3.62	19.71 (4.44)	N/A

[a]N/A: not applicable.

NPS

All participants completed the exit survey after taking the initial survey and being introduced to InHouse Call (Table 10). Of the 20 respondents to the exit survey, 95% (19/20) scored the likelihood of recommending InHouse Call to a friend as a ≥9 on a 10-point Likert scale, and were therefore classified as promoters. In total, 5% of the respondents scored this same question as 7 or 8 and were classified as neutral, while none provided a score of ≤7. This yielded an NPS of 95.

Table 10. Exit survey results.

Likeliness/usefulness rating[a]	Respondents' answers (N=20), n (%)			
	How useful did you find In-House Call?	How comfortable were you using InHouse Call?	How likely would you use InHouse Call in your daily work?	How likely would you recommend InHouse Call to another coworker?
1-7	0 (0)	0 (0)	0 (0)	0 (0)
8	1 (5)	2 (10)	3 (15)	1 (5)
9	2 (10)	4 (20)	1 (5)	2 (10)
10	17 (85)	14 (70)	16 (80)	17 (85)

[a]1=least likely, 10=most likely.

Feedback Results

Nearly all survey respondents participated in the optional write-in feedback section of the exit survey (Table 11). Participants identified three major categories of feedback on InHouse Call: (1) ease of use, (2) efficiency and usefulness in daily work, and (3) opportunities for improvement. Table 11 summarizes these categories and provides participant quotes for illustrative purposes.

Half of the participants commented on the ease of use, with some describing the interface as "user-friendly" and "intuitive." Nearly half of participants commented on the app's efficiency, with several mentioning the benefit of not having to wait on hold while making a telephone call.

Table 11. Exit survey with quotes.

Category of feedback	Example quotes
Ease of use	• "Love the easy access to all necessary #'s, esp RN pods." [PGY-4 Endocrinology Fellow] • "Awesome, easy to use, time saver, eliminates hassle of searching numbers." [Trauma Surgery advanced practice provider/nurse practitioner/physician assistant]
Efficiency and usefulness in daily work	• "App would be very useful." [Hospitalist] • "That is so much easier than using pocket cards or calling a main number to try to reach another department. This app would greatly improve my productivity." [PGY-3 Internal Medicine]
Opportunities for improvement	
Can provide even further available information, such as other departments and clinics	• "I would add charge nurse info in the room assignment search result. Make sure things like GI lab + pulm lab, etc." [General Surgery advanced practice provider/nurse practitioner/physician assistant]
Can add other information such as on call services and updated admission algorithms	• "ICU Attending #, VIR, CT Surgery, Off site surgeons (example: southern surgical)...agree with algorithm admissions, consult services." [PGY-3 Emergency Medicine]

Discussion

Principal Findings

In a series of trials with a variety of providers, mostly comprising internal medicine and family medicine residents and hospitalists but including surgical NPs or PAs and subspecialists, a total of 88 timed telephone calls were conducted as part of the timed trials to assess the effectiveness of InHouse Call. Traditional methods, such as using the operator services or pocket cards are cumbersome, antiquated methods for making calls in the hospital setting, and resulted in average trial times of 73.5 seconds per call.

By eliminating these time-consuming steps, the average time saved by using InHouse Call ranged from 45.71 seconds to 58.93 seconds. The time saved is significant when added over longer periods of time and with larger pools of hospital clinician users. Even when considering only the 10% of survey respondents who self-reported making 26-30 calls per day, the app would save those clinicians approximately 30 minutes of time spent on the phone per work day. This could be extrapolated by the estimated amount of phone calls made from providers to the hospital operator of 1000 calls per day, and we begin to see over 5000 work hours saved per year.

Wait times, transfers, and dropped calls were a major factor in the large degree of variance when using traditional methods to make a call. Conversely, InHouse Call eliminated these variable factors and streamlined the process, resulting in a more consistent outcome. This standardization in the process demonstrates the app's efficiency and was also reflected in the participants' feedback.

InHouse Call received an overwhelmingly positive response from its users with a strong NPS of 95, owing largely in part to the way it directly addressed participants' top complaints with the hospital communication system of "taking too much time" and "difficulty finding the right number." Eliminating long hold times and call transfers addressed the time-consuming

complaint, while the "Search Patient Room" database and subfolders addressed the difficulty in finding the correct contact. InHouse Call uses the patient's room number as an invariant through which a large amount of contact information can be quickly accessed. The subfolders are also effective at organizing contact information in large departments such as the radiology department with several modalities such as computed tomographic scan, magnetic resonance imaging (MRI), ultrasound, and ECHO available. The subfolders quickly and efficiently guide the user to the correct contact information compared to when using an often difficult to read pocket card.

Although the tallies for average time saved and an NPS of 95 are good prognostic indicators for the potential reliability of InHouse Call in a clinician's daily work, there are several future features which will further improve the app's utilization. Chief among them is the recently added ability to make calls directly from the app, thus further streamlining the process by eliminating the need to manually dial the number into the clinician's workstation phone. Furthermore, making the contact information in the subfolders (MRI, clinics, etc) part of the searchable database would further streamline the process of accessing that information. Survey respondents provided feedback for possible additional features including on-call providers and an admission algorithm, which may eliminate calls to the incorrect admission team.

In the exit survey, the lowest-scoring question was "How comfortable were you using InHouse Call?" despite feedback from the same survey respondents indicating that there was significant ease of use and an intuitive user interface. This may be accounted for by the fact that the respondents only had 45-60 seconds to familiarize themselves with the app before beginning their trials and filling out the exit survey. Their ease and comfort might be higher with a longer exposure time to the app and its functionality. A larger pilot study of tracking the app's usage in the clinician's workday would be advantageous to more accurately determine the app's integration into a provider's workflow. A system usability scale (SUS) would also be beneficial to quantifiably ascertain InHouse Call's usability in addition to the exit survey and feedback provided.

Comparison With Preexisting Systems

Improved hospital communication has been a focus of modern innovation for the past several decades. Several large medical centers have relied on direct messaging systems such as Cureatr, Cortext, and Voalte [13]. Although adequate in solving the problem of communication, the major limitation with these systems is that the name of the health care team member contact is required to initiate the communication; nonetheless, the team members' shifts change too frequently for this information to be kept up to date [14]. Further, these direct messaging systems only communicate with each other in a closed loop and thus add to an already complex hospital communication system. InHouse Call addresses these shortcomings by first having a database that is centered around a patient's room, and second by utilizing the already existing hospital telephone system with which users are familiar.

Some attempts have been made to create comprehensive hospital directories, but these systems have not reached mainstream success [15]. Their limitations include having their contact information crowd-sourced [16], or their systems were unable to alleviate frustration over finding the appropriate contact [17]. Studies regarding the usability and impact of other communication systems on clinical practice are limited; however, to our knowledge, this is the first study comparing timed trials of actual providers in their natural work environment.

Limitations

There are some limitations to this study. Since the app developer and author was the person who conducted the study, there is room for observation biases in the timed trials. However, the effectiveness of the app is self-evident in the exceptional time-saving results achieved. Further, the independent user feedback obtained in the exit strategy praised the app's time-saving functionality. Participant bias is also a risk owing to the participants being able to inherently see the aim of the study and thus put forth varying effort during the trials. Attempts to mitigate this bias were made by limiting the exposure of the participant to InHouse Call before the completion of the first round of the timed trials. In addition, the absence of incentives and rewards for the participants help to decrease bias with the results. Since the author and facilitator is also the app developer, there is an inherent conflict of interest not unique to studies of novel innovations. Conducting a larger study with more degrees of separation between the developer and participants would help mitigate this. Another limitation of the study was that it was performed at a single hospital. Although utilizing a phone operator system is common in all hospitals, there may be variables in different phone networks, which may impair InHouse Call's usability across all health care systems. However, InHouse Call is designed with the fundamental structure of a hospital in mind by utilizing the patient room number as the keystone of the contact database. Finally, owing to the busy work schedules of the participants and the somewhat lengthy study design, the number of volunteers, though diverse, was limited. Thus, the sample size was small. A larger cohort study would provide a more accurate insight into the app's receptivity.

Conclusions

We designed and implemented a novel native information directory app, InHouse Call, and found that its application to the average provider's workday saves a significant amount of time in placing calls. By eliminating wait times, call transfers, and dropped calls, the average amount of time to initiate and complete a call was significantly reduced as well as the variability of time to complete the task. Users found the app easy to use, effective, and useful for their daily work. The NPS was an astounding 95, which is on par with other great apps. Despite its current effectiveness, opportunities for improvement were also determined. As InHouse Call relies on the current telephone system already universally found in large hospitals, it has the potential to be expanded to nearly all other institutions.

Conflicts of Interest

GS is the developer of InHouse Call.

References

1. Reynard J, Stevenson P. Practical Patient Safety. Oxford: Oxford University Press; 2009.
2. Sijstermans R, Jaspers MWM, Bloemendaal PM, Schoonderwaldt EM. Training inter-physician communication using the Dynamic Patient Simulator. Int J Med Inform 2007;76(5-6):336-343. [doi: 10.1016/j.ijmedinf.2007.01.007] [Medline: 17331800]
3. Richardson JE, Ash JS. The effects of hands-free communication device systems: communication changes in hospital organizations. J Am Med Inform Assoc 2010;17(1):91-98 [FREE Full text] [doi: 10.1197/jamia.M3307] [Medline: 20064808]
4. Wu RC, Tran K, Lo V, O'Leary KJ, Morra D, Quan SD, et al. Effects of clinical communication interventions in hospitals: a systematic review of information and communication technology adoptions for improved communication between clinicians. Int J Med Inform 2012 Nov;81(11):723-732. [doi: 10.1016/j.ijmedinf.2012.05.014] [Medline: 22727613]
5. Agarwal R, Sands DZ, Schneider JD. Quantifying the economic impact of communication inefficiencies in U.S. hospitals. J Healthc Manag 2010;55(4):265-81; discussion 281. [Medline: 20812527]
6. Little AD. Telecommunications: can it help solve America's health care problems?. Cambridge, MA: Arthur D Little; 1992.
7. How house officers spend their nights. N Engl J Med 1989 Dec 07;321(23):1615-1616. [doi: 10.1056/nejm198912073212317] [Medline: 2586563]
8. Parenti C, Lurie N. Are things different in the light of day? A time study of internal medicine house staff days. Am J Med 1993 Jun;94(6):654-658. [doi: 10.1016/0002-9343(93)90220-j] [Medline: 8506894]
9. Block L, Habicht R, Wu AW, Desai SV, Wang K, Silva KN, et al. In the wake of the 2003 and 2011 duty hours regulations, how do internal medicine interns spend their time? J Gen Intern Med 2013 Aug;28(8):1042-1047 [FREE Full text] [doi: 10.1007/s11606-013-2376-6] [Medline: 23595927]
10. Coiera E. Medical informatics. BMJ 1995 May 27;310(6991):1381-1387 [FREE Full text] [doi: 10.1136/bmj.310.6991.1381] [Medline: 7787545]
11. Reichheld FF. The One Number You Need to Grow. Harvard Business Review. 2003. URL: https://hbr.org/2003/12/the-one-number-you-need-to-grow [accessed 2018-03-02]
12. Schumacher RM, Lowry SZ. NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records. National Institute of Standards and Technology. 2010. URL: https://www.nist.gov/system/files/documents/itl/hit/Guide_Final_Publication_Version.pdf [accessed 2018-03-21]
13. Brusco JM. Using smartphone applications in perioperative practice. AORN J 2010 Nov;92(5):503-508. [doi: 10.1016/j.aorn.2010.09.001] [Medline: 21040814]
14. Przybylo JA, Wang A, Loftus P, Evans KH, Chu I, Shieh L. Smarter hospital communication: secure smartphone text messaging improves provider satisfaction and perception of efficacy, workflow. J Hosp Med 2014 Sep;9(9):573-578 [FREE Full text] [doi: 10.1002/jhm.2228] [Medline: 25110991]
15. DirectLineMD. URL: https://www.directlinemd.com/ [accessed 2022-01-19]
16. Induction Switch. Induction Healthcare. URL: https://induction-app.com [accessed 2022-01-19]
17. Morawski K, Monsen C, Takhar S, Landman A. A Novel Information Retrieval Tool to Find Hospital Care Team Members: Development and Usability Study. JMIR Hum Factors 2018 Apr 16;5(2):e14 [FREE Full text] [doi: 10.2196/humanfactors.6781] [Medline: 29661752]

Abbreviations

APP: advanced practice provider/nurse practitioner/physician assistant
ECHO: echocardiogram
HIPAA: Health Insurance Portability and Accountability Act
MRI: magnetic resonance imaging
NP: nurse practitioner
NPS: net promoter score
PA: physician assistant
RN: registered nurse

Edited by A Kushniruk; submitted 28.06.21; peer-reviewed by A Islam, L Guo, Y Tani; comments to author 13.10.21; revised version received 04.11.21; accepted 27.11.21; published 27.01.22.

Please cite as:
Schilling G, Villarosa L
An Information Directory App InHouse Call for Streamlining Communication to Optimize Efficiency and Patient Care in a Hospital: Pilot Mixed Methods Design and Utility Study
JMIR Hum Factors 2022;9(1):e23833
URL: https://humanfactors.jmir.org/2022/1/e23833
doi:10.2196/23833
PMID:35084350

Original Paper

Enriching the Value of Patient Experience Feedback: Web-Based Dashboard Development Using Co-design and Heuristic Evaluation

Mustafa Khanbhai[1], BSc, MBChB, PhD; Joshua Symons[1], CDMP, MSc; Kelsey Flott[1], BA, MSc, PhD; Stephanie Harrison-White[2], BSc; Jamie Spofforth[2], MSc; Robert Klaber[2], MA, MBBS, MD; David Manton[1], BSc, MSc, PhD; Ara Darzi[1], KBE; Erik Mayer[1], MBBS, PhD

[1]Patient Safety Translational Research Centre, Imperial College London, National Institute for Health Research/Institute of Global Health Innovation, London, United Kingdom

[2]Imperial College Healthcare NHS Trust, London, United Kingdom

Corresponding Author:
Mustafa Khanbhai, BSc, MBChB, PhD
Patient Safety Translational Research Centre
Imperial College London
National Institute for Health Research/Institute of Global Health Innovation
Queen Elizabeth the Queen Mother Wing
St Mary's Campus
London, W2 1NY
United Kingdom
Phone: 44 02033126428
Email: m.khanbhai@imperial.ac.uk

Abstract

Background: There is an abundance of patient experience data held within health care organizations, but stakeholders and staff are often unable to use the output in a meaningful and timely way to improve care delivery. Dashboards, which use visualized data to summarize key patient experience feedback, have the potential to address these issues.

Objective: The aim of this study is to develop a patient experience dashboard with an emphasis on Friends and Family Test (FFT) reporting, as per the national policy drive.

Methods: A 2-stage approach was used—participatory co-design involving 20 co-designers to develop a dashboard prototype, followed by iterative dashboard testing. Language analysis was performed on free-text patient experience data from the FFT, and the themes and sentiments generated were used to populate the dashboard with associated FFT metrics. Heuristic evaluation and usability testing were conducted to refine the dashboard and assess user satisfaction using the system usability score.

Results: The qualitative analysis from the co-design process informed the development of the dashboard prototype with key dashboard requirements and a significant preference for bubble chart display. The heuristic evaluation revealed that most cumulative scores had no usability problems (18/20, 90%), had cosmetic problems only (7/20, 35%), or had minor usability problems (5/20, 25%). The mean System Usability Scale score was 89.7 (SD 7.9), suggesting an excellent rating.

Conclusions: The growing capacity to collect and process patient experience data suggests that data visualization will be increasingly important in turning feedback into improvements to care. Through heuristic usability, we demonstrated that very large FFT data can be presented in a thematically driven, simple visual display without the loss of the nuances and still allow for the exploration of the original free-text comments. This study establishes guidance for optimizing the design of patient experience dashboards that health care providers find meaningful, which in turn drives patient-centered care.

(JMIR Hum Factors 2022;9(1):e27887) doi:10.2196/27887

KEYWORDS

patient experience; friends and family test; quality dashboard; co-design; heuristic evaluation; usability

Introduction

Patient Experience

Understanding patients' experience of health care is central to the process of providing care and is a fundamental pillar of health care quality. It is now widely acknowledged that patients want to give feedback about health care [1] and that staff should be listening to what their patients say about the experience of being in the hospital. However, whether staff can use this feedback to make changes to improve patients' experiences is now a national initiative [2-6]. This pertains to differing areas of the health care system, from senior management at the level of the hospital board down to individual frontline health care staff. There is a concern that the ever-growing collection of feedback is not being used for improvement but rather represents a *tick box mentality* of organizations thinking that they are listening to their patients' views but not actually doing so [7]. Several studies have looked at teams of frontline staff to understand how ward staff can engage with patient feedback to make meaningful improvements [1,4,5,8]. Most of the literature in this area finds that despite enthusiasm to make improvements and the vast rhetoric around this, proactive changes are often minimal and largely concentrated on "quick fixes" [3].

Using Patient Experience Data to Drive Change

Health care organizations within the English National Health Service (NHS) have received recent encouragement to understand the ways in which they use patient feedback to improve care [9]. NHS England and NHS Improvement have implemented changes in patient experience data collected via the Friends and Family Test (FFT). One area of focus is placing greater emphasis on the use of FFT data to drive improvement. For health care organizations to act on this policy change, they need to tackle both macrolevel factors (how organizational structures are unwittingly preventing progress) and microlevel factors (how individual clinicians and teams of staff have difficulty engaging with the data sources) [4]. An organizational strategic focus that prioritizes use over collection and ensures data are relayed to staff by patient experience teams in an accessible, straightforward, and engaging manner is required. Staff training on both quantitative and qualitative analytical techniques and quality improvement (QI) methodologies is also needed. There should be an organizational emphasis where patient experience data collected can be meaningfully used by frontline staff.

Visualizing Patient Experience Data Through Web-Based Dashboards

There is some evidence that implementing quality dashboards provides constant access to information that can improve adherence to quality guidelines and may help improve patient outcomes [10]. Key reports have called for comprehensive, real-time health care information technology to be integrated into clinical and management processes in health care to improve quality and patient safety [11-13]. A recent report by the National Institute for Health Research [14] recommends that health care organizations produce dashboards and describes dashboards as essential tools to help staff understand areas for improvement in a timelier manner. Visualization of patient feedback is crucial for helping frontline staff and key stakeholders make sense of the structure and underlying patterns in their patients' experiences. The insights gained from these underlying patterns have the potential to answer vital questions at the point of care [15]. To facilitate this, engaging staff and patients using a co-design approach to visualize feedback is likely to result in sustainable improvements at a local level. Co-design is a process in which targeted end users and other relevant stakeholders form a partnership with researchers and work together on all aspects of intervention development, from needs assessment to content development, pilot-testing, and dissemination [16]. Co-designed interventions may be more effective than traditional approaches where interventions are largely designed by researchers and clinicians. This approach increases the involvement of key stakeholders by encouraging a bottom-up approach, thereby helping health care organizations think differently [17]. The aim of this study is to develop a patient experience dashboard with an emphasis on FFT reporting, as per the national policy drive. An iterative process involving co-design with key stakeholders was used to develop the dashboard, followed by heuristic usability testing.

Methods

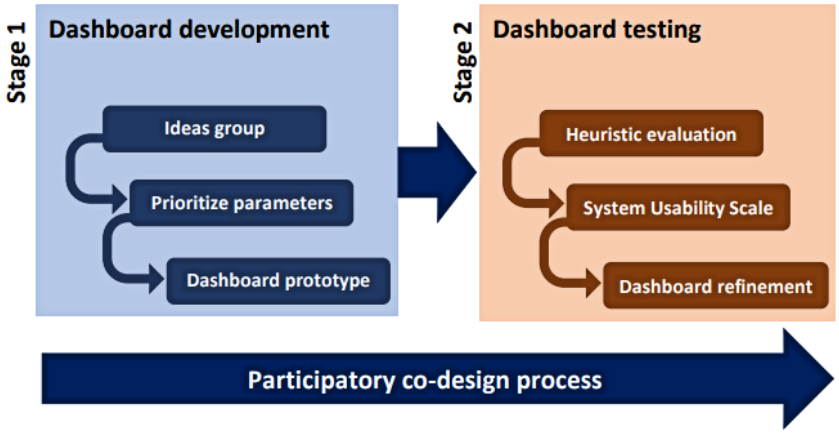
Setting

This study was conducted at a large London NHS trust. Alongside other health and care partners, the trust caters to a population of approximately 2.3 million people across its 5 sites. Services include accident and emergency, inpatient, outpatient, and maternity, which routinely collect FFT patient experience data. At least one stakeholder from each of the 4 service settings participated in this study.

Study Design

This study had two key design stages: (1) dashboard development and (2) dashboard testing (Figure 1). Stage 1 followed a participatory co-design process, which involved key stakeholders (namely health care staff, managers, and a patient representative) in the design of the dashboard. Stage 2 involved heuristic assessment to conduct an informal usability inspection of the dashboard to evaluate whether the user interface of a system adhered to a set of usability principles known as heuristics [18,19]. An invitation letter and a participant information sheet were emailed to the key stakeholders. Informed consent was obtained before interview participation. The researchers led the ideas groups, facilitated and summarized the discussions, took field notes, and made audio recordings. The study received ethical approval from North East–Tyne and Wear South Research Ethics Committee, 17/NE/0306, and took place between April 2018 and February 2019.

Figure 1. Participatory co-design process used in the study, including stage 1 (developing the dashboard) and stage 2 (testing the dashboard).



Patient Experience Reports

Retrospective FFT data were used for the intervention from January 2017 to July 2017. We extracted approximately 69,285 FFT responses and associated comments, which were considered adequate to demonstrate the satisfactory accuracy of the text analytics software [20]. Free-text fields identifying favorable service (*What did we do well?*) and areas requiring improvement (*What could we do better?*) were extracted from the patient experience reports across 4 care settings. Using language analysis, free-text comments were themed according to the 2011 English NHS Patient Experience Framework [21] (Textbox 1). The framework was developed by the NHS National Quality Board to guide the measurement of patient experience across the NHS. The framework outlines the elements that are critical to the patients' experience of NHS services. Sentiment analysis was performed for each free-text comment (ie, positive, negative, or equivocal; Textbox 2). The free-text data (themes and sentiments) from the language analysis and associated FFT parameters were presented to the key stakeholders to develop a bespoke dashboard.

Textbox 1. The eight themes that outline the elements that are critical to patients' experiences [21]. The themes in italics were added by the research team to include comments that did not fall into the original framework themes.

Patient Experience Framework themes

- Respect for patient-centered values, preferences, and expressed needs
- Coordination and integration of care
- Information, communication, and education
- Physical comfort
- Emotional support
- Welcoming the involvement of family and friends
- Transition and continuity
- Access to care
- Staff
- General
- Unclassified

Textbox 2. The four sentiment categories used to classify patient experience Friends and Family Test free-text comments.

Sentiment categories

- Positive
- Neutral
- Negative
- Unclassified

Participatory Co-design

In the co-design process, the concepts *Ideas groups*, *Stakeholder needs*, and *Prototyping* were used as described in the Health Service Co-design toolkit [22]. This meant having an iterative refinement process that was reactive to the participants' feedback. Ideas groups are a tool used to brainstorm ideas for improvement and ways of implementing them in clinical practice. A stakeholder needs table was a useful tool for sketching out possible improvements near the start of the co-design as well as deciding on key areas for improvement and specific improvements later. Prototyping was used to test new products to see if they worked and was a useful way to engage and stimulate creativity among the stakeholders taking part in ideas groups [22].

User Needs for Patient Experience (FFT) Reporting

Through purposive sampling, we began by identifying staff to act as co-designers within the patient experience team, followed by lead nurses and junior staff based in 4 care settings: inpatient, outpatient, accident and emergency, and maternity. This strategy helped ensure that we included staff that were either directly or indirectly involved in patient care. A criterion required the interviewees to have a good overview of patient experience feedback, including the FFT, and be currently using all or part of this service in their everyday activities.

Stakeholders came together in ideas groups, and the aim was described: to develop a dashboard that would allow for succinct visualization of the analyzed free text from the FFT reports to identify areas of improvement in a timely manner. In the first ideas group, we discussed the parameters from the FFT reports, including the free-text language analysis output ideas (Textbox 3) and key requirements of the dashboard that were deemed important. In total, 2 research facilitators (MK and SHW) began each 90-minute focus group with a brief presentation of the language analysis toolkit followed by a display of all the FFT parameters, including the free-text language analysis output, separated by theme and sentiment. The participants then broke into 5 groups of 4 to brainstorm ideas on how the FFT reports (ie, core parameters and key requirements that needed to be incorporated into the dashboard) and the various display formats were sketched, and the reasons were presented by a representative in each group. Each participant then had to independently rank their preferred display format (ie, rank order from 1=*first* to 4=*last*). The interviews were transcribed verbatim and double-checked for inaccuracies. To aid trustworthiness of data collection, the first author checked accuracy against interview audio recordings, and the participants were asked to review the transcript of their interview and any sensitive comments were redacted before analysis.

Textbox 3. Friends and Family Test (FFT) questions, including supplementary questions and associated parameters that are routinely collected as part of the FFT survey.

FFT questions

- How likely are you to recommend our service to friends and family if they needed similar care or treatment?
  - Extremely likely
  - Likely
  - Neither likely nor unlikely
  - Unlikely
  - Extremely unlikely
  - Don't know
- What did we do well?
  - Patient experience theme
  - Sentiment
- What could we do better?
  - Patient experience theme
  - Sentiment

Associated parameters routinely collected as part of the FFT survey

- Date
- Hospital
- Division
- Ward or clinic
- Language used
- Channel
- Responder (patient, carer, or family)
- Gender
- Ethnicity
- Age range
- Disability

Development of the Prototype Dashboard

A prototype was developed and sent out to all the participants. Within 2 weeks, a member of the research team (MK) visited each participant to understand stakeholder needs and gather information on the prototype design (including layout, colors, and information presented on the dashboard) and suggestions for improvements. In general, these feedback sessions lasted approximately 10 to 15 minutes.

Heuristic Evaluation and Usability Testing

The primary goal of this evaluation was to reduce errors in interpretation and accommodate rapid comprehension, which is critical for using FFT reports in a timely manner. This heuristic evaluation was our initial step toward the development of FFT-visualization-specific heuristics. For this study, we used a validated heuristic evaluation checklist developed to evaluate systems that produce information visualizations [23]. The principles from the heuristics by Nielsen [24] were combined with heuristic principles developed specifically to evaluate information visualization. The use of evaluators who are experts in visual design and understand the analytic intent of the visualizations was important. This was conducted by JS, who has health and design expertise, and by RK and MU, who have health, design, and QI expertise. The checklist consists of 10 usability principles substantiated with 49 usability factors. If the factor was present, the evaluator gave a score of 1 (*Yes*) and, if it was not present, they gave a score of 0 (*No* or *N/A*) [25]. The evaluators drew from heuristic principles related to visual and graphical perception and best practices in graph design as well as years of experience in clinical practice and QI.

The System Usability Scale (SUS) [26], which is a validated posttest questionnaire, was used to measure user satisfaction with product usability. It consists of 10 statements that are scored on a 5-point scale of strength of agreement that captures ratings of electronic devices or systems, including respondent assessments of future use, complexity, ease of use, and perceived

usefulness of the display of results. The questionnaire provides a score (range 0-100) based on a participant's rating of 10 statements regarding a product's usability. Higher scores indicate greater satisfaction with usability. As a general rule, a system with a score of >70 has acceptable usability; a lower score means that the system needs more scrutiny and continued improvement [27].

Data Analysis

Data from the ideas groups and from the open-ended questions in the questionnaires were evaluated, discussed, and summarized by the research group. As the aim was to identify improvement ideas expressed by the participants and evaluate the intervention, the data were summarized without an in-depth qualitative analysis. Descriptive statistics were used to describe the participants' background characteristics. Frequencies and proportions were used to describe the outcomes of the questionnaires and were calculated using Microsoft Excel (version 2019).

The SUS was scored by converting responses to a 0-4 scale (4 was the most positive response). The converted responses were added and multiplied by 2.5, as per the scoring instructions, giving a range of possible values from 0 to 100. Descriptive statistics were used to summarize the SUS scores across all the evaluators of the system. The output from a heuristic evaluation is a summary list of usability problems identified by the group of evaluators. The scores for each heuristic were calculated by dividing the total number of factors (points) awarded by the total number available. The higher the score, the more usable the system was considered to be.

Results

Co-designers' Characteristics

A total of 20 co-designers were recruited for this project (Table 1). We selected co-designers with a variety of characteristics in terms of their professional background, the service settings (division) they were employed in, and whether they were clinical or nonclinical to ensure that the development of the dashboard took into account a diversity of participants.

Table 1. Characteristics of the co-designers (N=20).

Characteristic	Participants, n (%)
Professional background	
Nursing and midwifery	6 (30)
Allied health	2 (10)
Medical	2 (10)
Nonclinical service	
Patient experience	3 (15)
Quality improvement	3 (15)
Data analytics	2 (10)
Health care design	2 (10)
Division	
Surgery and cancer	3 (15)
Medicine and integrated care	4 (20)
Women and children, and clinical support	3 (15)
Nonclinical service	
Patient experience	3 (15)
Quality improvement	3 (15)
Data analytics	2 (10)
Health care design	2 (10)

Participatory Co-design Process

The participants were generally enthusiastic about the development of a visualization tool for displaying FFT data and, in particular, the free-text comments in a meaningful way and in near real time. Most felt that a dashboard might highlight areas that required improvement as well as areas that had been improved, which might enhance how staff interacted with FFT data. Results from the ideas group were separated into the FFT parameters that were deemed important, key requirements that should be considered during development, and ranking of the 4 dashboard sketches.

FFT Parameters

The feedback from the ideas group highlighted that, although all parameters were important, only a select number were chosen to be displayed on the opening screen, whereas the rest could be accessed through a tab. The most important parameters were date, ward or division, sentiment, and patient experience theme. As the FFT is anonymous, most staff thought that segmenting

the feedback by demographics had a risk of identifying the patient, especially if the reports were accessed in real time. The date of feedback was crucial to respond in near real time and to look for trends and assess progress over time. The ward or division was required so that feedback could be accessed by all staff, ensuring transparency as well as identifying opportunities for improvement (eg, from other wards with similar specialties or patient profiles). The FFT score was considered less insightful in understanding where improvements needed to be made, and the participants unanimously agreed that the free-text option should take precedence when displaying the FFT data and should be displayed on the opening dashboard screen. Individual sentiment was not considered useful as most were positive; however, the average sentiment of each patient experience theme was the preferred approach. The participants highlighted that negative comments could be sandwiched between positive comments and vice versa and that staff felt it was important to consider this context rather than separate the positive comments from the negative comments. Therefore, the themes with average sentiment were displayed as *to improve* in relation to the question *What could we do better* and as *doing well* in relation to the question *What did we do well?* Despite the free-text comments being clustered into themes, frontline staff agreed that they should have the opportunity to drill down into specific or unusual comments for further manual analysis to gain additional insight.

Key Dashboard Requirements

We summarized feedback from the ideas group on what an ideal dashboard would require in relation to FFT reporting (Textbox 4). The statements reported related to accessing the reports in an easy and understandable manner that allowed staff to assimilate the pertinent information in a short time frame, thereby addressing patients' experiences as they are reported.

Textbox 4. A summary of the key requirements for the dashboard from the ideas group.

Key requirements for the dashboard

- Easy access to the data in a visual and usable format
- Data provided in a way that can be engaged with by frontline staff
- Summary data that can be mined down to individual comments
- Locally relevant information displayed for comparison across similar wards
- Ability to see change through the months or years
- Facilitating discussion with the executive board acting as leverage to drive change
- Information provision in near real time
- Positive feedback, celebratory sharing with teams
- Free text better than scores
- Giving all ward staff ownership of the data, narrowing the skill gap
- Content should not be overwhelming
- Imparting a positive mindset to improvement as core activity

Dashboard Design Popularity

A total of 4 main dashboard design formats were presented by the 4 groups: bar chart, line graph, bubble chart, and pictograph. Table 2 shows the preference rankings. The bubble chart was ranked first, being the most preferred by the participants (P<.001). This was primarily because the participants favored displaying the experience visualizations using the same format as other visualizations currently used in the organization, for example, the Patient Safety dashboard. This consists of the safety incidents using a bubble chart, which is currently used by all staff within the organization.

Table 2. Mean preference ranking (1=lowest and 4=highest) for each display dashboard among the co-design participants (N=20).

Dashboard design format	Preference ranking, mean (SD)
Bar chart	3 (0.86)
Line graph	1.35 (0.59)
Bubble chart	3.5 (0.69)
Pictograph	2.05 (0.89)

Development of the Dashboard Prototype

On the basis of observations, interviews, and feedback, we developed an information-rich suite of display implemented in Tableau (Tableau Software) that provided at-a-glance information of FFT-reported free-text data. Tabs for each dashboard were visible across all views that document the individual steps taken to develop the final dashboard. However, for the dashboard testing, the dashboard was presented on a Tableau reader, which does not allow the user to make any changes. The census overview was the opening screen, which contained the top 5 themes with the most negative sentiment presented on the left as *to improve* and the top 5 themes with positive sentiment presented on the right as *doing well* for all inpatient comments (Figure 2). A traffic light color coding system was developed (ie, the most negative sentiment was coded as red and the most positive sentiment was coded as green presented as a word heat map; Figure 2). The user had the ability to configure their preferences by isolating the visualization based on positive or negative sentiment instead of side-by-side comparison (Figure 3). There was a date range toggle bar and a list of wards on the right side of the dashboard screen that could be selected by the user or where the number of comments and average sentiment in each theme *bubble* could be viewed by hovering over each data point. The final version embodies a dashboard where users can interact with the visualization, use filters to modify the display, and select an individual theme *bubble* that presents all the free-text comments within that specific theme (Figure 3).

Figure 2. Prototype dashboard presented in a bubble chart, where inpatient free-text comments are split by the top 5 themes and sentiment (negative [to improve] on the left and positive [doing well] on the right). A word heat map demonstrates the most common words found within the free-text comments, followed by individual comments.

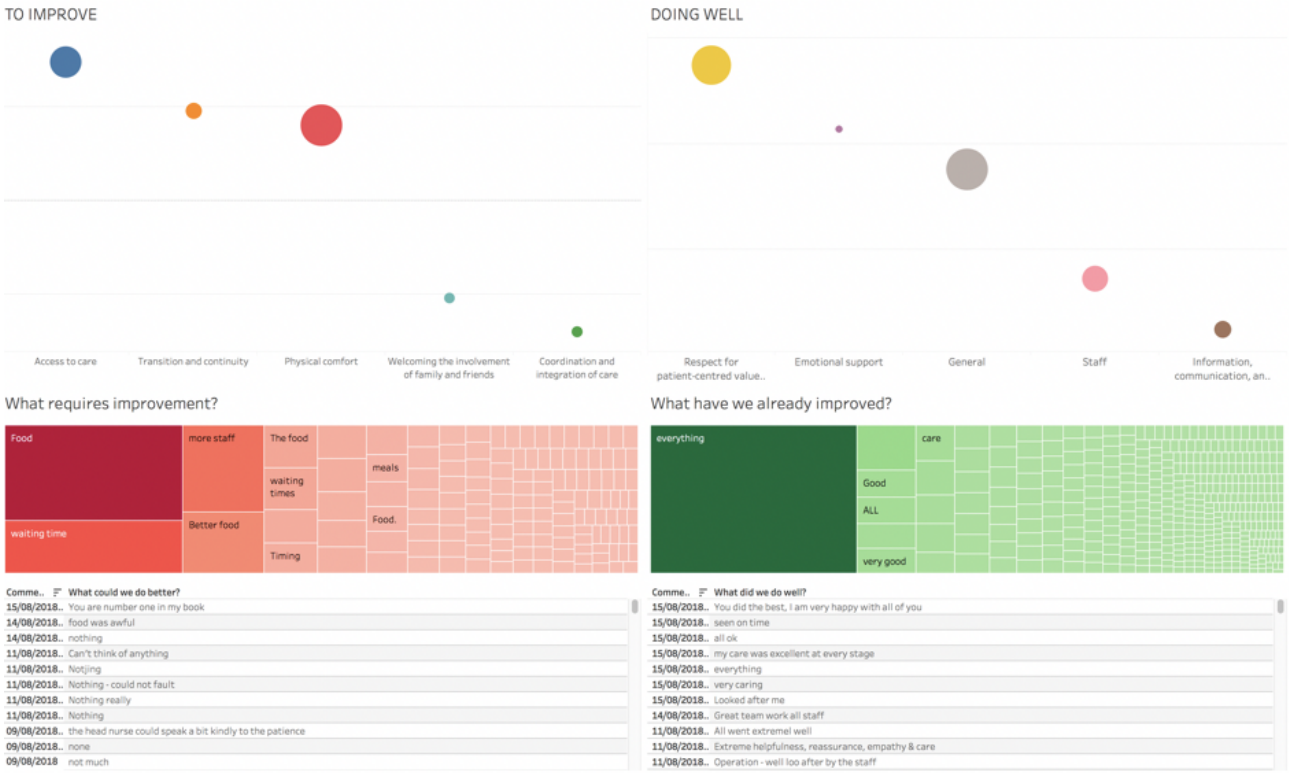
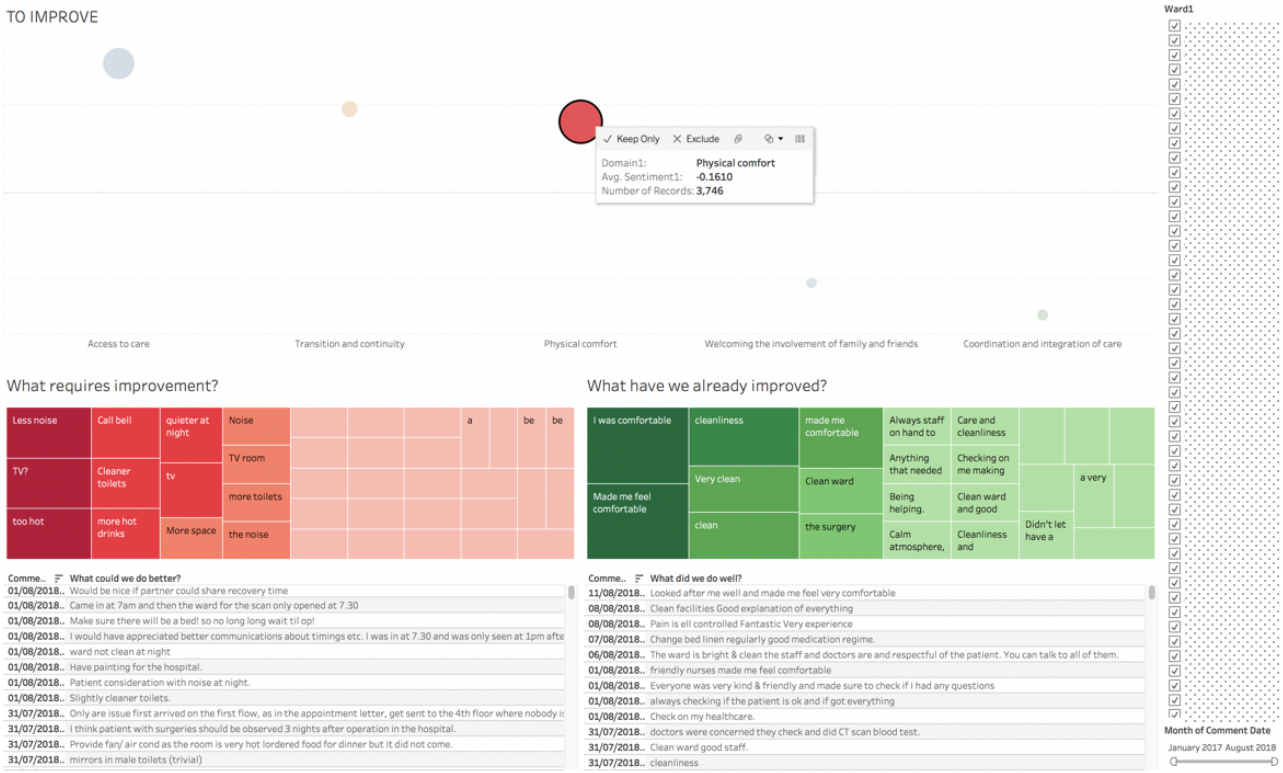


Figure 3. This display demonstrates only negative (to improve) inpatient comments with web-based features. The word heat map shows the most common comments split by negative sentiment in red and positive sentiment in green, followed by individual comments that describe physical comfort only.



Heuristic Evaluation and Usability Testing

Most cumulative scores from the 3 participants who took part in the heuristic evaluation had no usability problems (18/20, 90%), had cosmetic problems only (7/20, 35%), or had minor usability problems (5/20, 25%). The areas requiring attention recorded by a higher severity rating were user control and freedom, and consistency and standards. The percentage score was lowest in user control and function (60%) followed by consistency and standards (66.7%), and the highest score was flexibility and efficiency of use (90%) followed by visibility of system status (88.3%; Table 3). The heuristic evaluators also made suggestions for their implementation. Specific issues that required addressing were having the dashboard service settings consistent (eg, inpatient compared with maternity; minor usability problem), making the data accessible on hovering the mouse (cosmetic problem only), ensuring the data were presented as the 2 supplementary questions (cosmetic problem only), changing all font to *Arial* (cosmetic problem only), increasing the size of the *bubble* (cosmetic problem only), presenting the data in descending order and having the month toggle bar at the top of the screen (minor usability problem), and excluding the comments themed as general (minor usability problem). There was unanimous feedback that the word-based heat map, although useful, did not add much to gaining knowledge and made the dashboard cluttered (minor usability problem); however, the color coding should remain for the headings (ie, green for positive sentiment and red for negative sentiment) and the caption above the comments should be removed (cosmetic problem only). As the dashboard was presented on a Tableau reader that did not allow the participants to make any changes, some of the questions about user control and freedom did not apply; however, the free-text responses from the participants were taken into account.

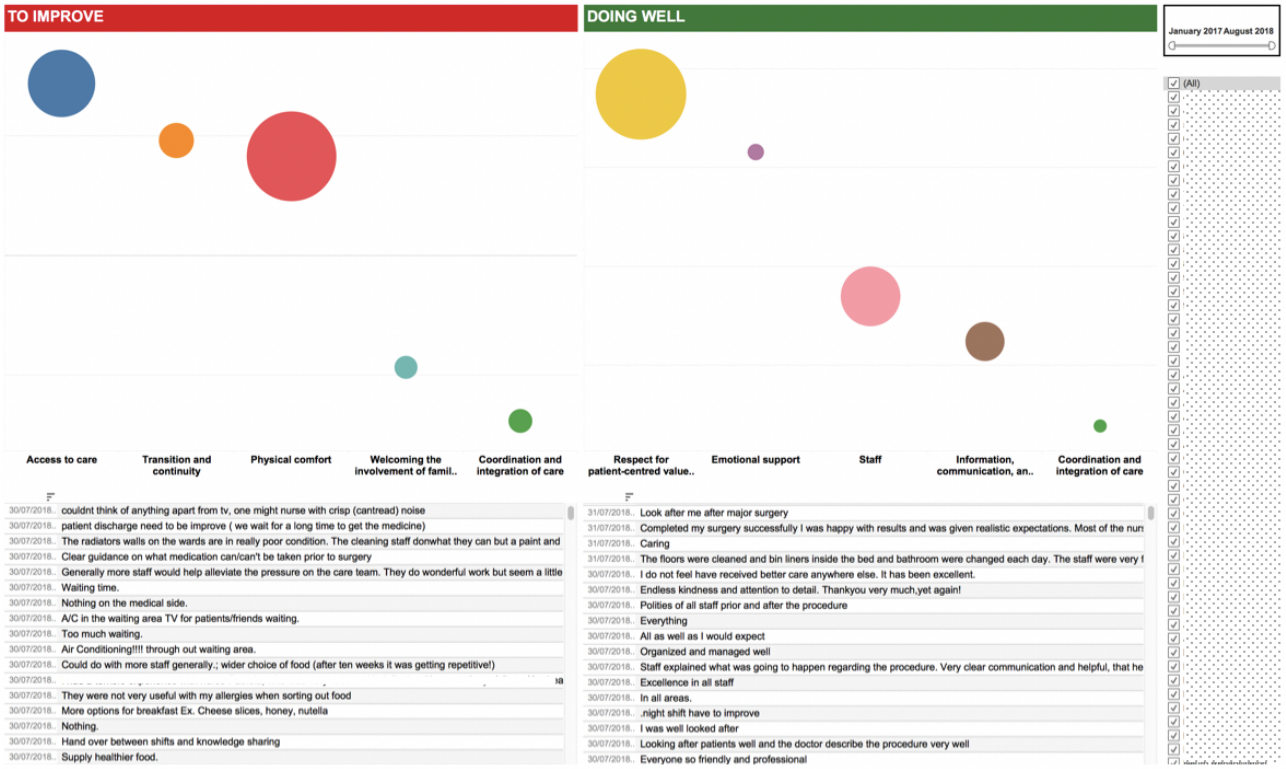
Amendments were made accordingly, and the final dashboard (Figure 4) was tested on the participants for satisfaction. The mean SUS score was 86.97 (SD 5.79), and the median score was 87.5. Participants from a nursing background and those from the patient experience team with a nonclinical background had the highest scores.

Table 3. Mean heuristic evaluation ratings for the prototype dashboard.[a]

Heuristic evaluation (maximum score)	Overall severity rating, mean (SD)	Score, mean (SD)	Score result (%)
Visibility of system status (6)	0 (0)	5 (1)	83.3
Match between system and the real world (5)	0.7 (1.2)	4 (1)	80
User control and freedom (5)	1.7 (0.6)	3 (1)	60
Consistency and standards (6)	1.3 (0.6)	4 (1)	66.7
Recognition rather than recall (4)	0.7 (1.2)	3 (1)	75
Flexibility and efficiency of use (7)	0.7 (0.6)	6.3 (0.6)	90
Esthetic and minimalist design (7)	0 (0)	5 (1)	71.4
Spatial organization (3)	0 (0)	2.3 (0.6)	77.8
Information coding (2)	0.7 (0.6)	1.3 (0.6)	66.7
Orientation (4)	0 (0)	3 (0)	75

[a]The maximum score that each question can receive is shown in parentheses. The overall severity rating score ranges from 0 (no usability problem) to 4 (usability catastrophe), and the mean overall severity rating is shown. The score result is calculated as a percentage of the maximum score.

Figure 4. Final dashboard amended following heuristic evaluation, which was tested using the System Usability Scale. This dashboard presents inpatient comments divided into the top 5 themes in descending order with negative (to improve) and positive (doing well) sentiment.



Discussion

Principal Findings

Quality dashboards integrating health care data offer innovative means of providing metrics that can facilitate QI [28]. We demonstrated an iterative approach to developing a web-based dashboard using free-text FFT data collected as part of a national drive to improve quality. To the best of our knowledge, this is the first study using co-design principles and heuristic evaluation to develop a dashboard to visualize free-text FFT data for QI.

The literature suggests that data availability is a crucial precondition for the development of dashboards [29]. However, this presents health care organizations with a challenge as data are often presented in quantitative and summative format, whereas staff also desire qualitative information [29,30], confirming our findings. Therefore, facilitated by the findings from our co-design study, we extended the scope of the FFT data by augmenting the dashboard with associated free-text data, which not only provides a richer narrative but also makes the data more meaningful to staff [5].

Previous research shows that actual dashboard development often starts with the translation of available data into useful dashboard content [29], and the use of focus groups facilitates a better understanding of the needs and wishes of the

stakeholders when formulating a design [28]. Our focus groups were guided by co-design principles—involving stakeholders in the design and development of visualization tools increases functionality and usability by meeting stakeholder requirements, thereby improving the quality of the system and increasing the likelihood of achieving intended health outcomes [31]. The literature on dashboard development mostly addresses the technical aspects of development processes while overlooking the organizational aspects [28]. Combining technical and organizational aspects into one comprehensive development process was vital for addressing this challenge. Our 2-stage approach illustrates a range of stakeholder engagement methods, dashboard prototypes, and design insights on meaningful dashboard content, format, and clinical use.

A stark finding from Weggelaar-Jansen et al [28] revealed that no hospital taught health care professionals or managers to understand statistical measurements and the related graphics to help them understand the dashboard. In addition, the studies [32,33] that used co-design to create a dashboard for Cancer Patient Experience Survey data and the patient experience toolkit did not clearly address the time poverty that is a growing challenge, hindering health care staff from having dedicated time within their duties to engage with the dashboard. Addressing these issues was a key requirement that was raised during the ideas group discussion in our study. The participants raised the issue that the pre-existing format of patient experience reporting used too much technical language, which required training in data analysis and statistics to facilitate its full understanding to then use the results appropriately. This aspect was particularly important to ensure that the dashboard was interwoven into the daily activities of frontline staff. Therefore, the participants unanimously agreed to create a patient experience dashboard that would follow an existing format that was established and widely used in the organization—the Patient Safety dashboard. This meant that the prototype evolved and adapted exploring similarities with the Patient Safety dashboard but displayed patient experience data that would enable staff to meaningfully engage with the new dashboard without *costing* so much of their time. Adopting a new visualization that was different from the format of the currently used Patient Safety dashboard would have resulted in a steep learning curve and possibly discouraged and disengaged staff, thereby failing to translate FFT reports into actionable interventions.

To achieve a broadly comprehensible layout, we ensured that the real-time graphic and visual presentation of the content fit the purpose of the dashboard [10]. Previous studies [32,33] have highlighted the use of visual and physical media as a form of sharing and communicating, which helped remove barriers to mutual understanding. Short summaries (eg, dashboards and graphs) are essential tools to help staff understand areas for improvements quickly [32] as the presentation of data enables them to navigate it in ways that answer questions specific to their service or to particular patients [33]. Embedding the outcomes of the participatory co-design process informed the development of the prototype, and validation with stakeholders using established usability techniques provided reassurance that the approach had value for staff.

To enable stakeholders to customize dashboard content to their own needs, research suggests that health care organizations add 3 main functionalities, namely, drill-down, filter, and alert functions [28]. Our dashboard fulfilled these criteria with the availability of filters to modify the display and select an individual theme *bubble* to present all the free-text comments within that specific theme and sentiment, and the ability to view the 5 most important themes as determined by sentiment. An interesting trade-off was observed between the need for detail and the need for brevity during the usability evaluation. Feedback from the heuristic evaluation demonstrated that the appearance of the dashboard needed to be simple and that it should not look like a major task to understand the features. Through a series of adaptations, we addressed the cosmetic problems (n=7) and minor usability problems (n=5) to deliver a punchy dashboard and still contain all the desired features and requirements that had been highlighted during the co-design process. These dashboard features specifically improved staff engagement and empowerment by attracting their attention and stimulating them to pay attention to the information of interest, keeping their attention and interest for longer periods, and providing a greater depth of content [19]. This meant that the final dashboard was ultimately designed for use by all health care staff, as demonstrated by the usability score. In general, it is considered that usable products should have SUS scores of >70; our prototype had a mean SUS score of 86.97 (SD 5.79), suggesting acceptable usability [27]. The highest scores came from the participants from a nursing background and the patient experience team, which is an encouraging result. We hope that this translates to sustained engagement in the use of the dashboard and, as a consequence, generates a body of *patient experience ambassadors* to help raise awareness of the use and importance of patient experience dashboards across the organization.

Limitations

This study was conducted at a single hospital, and all participants were employees within the hospital, thereby causing selection bias. Although this is a limitation, the principles underlying the development of the dashboard are transferable across different hospitals that collect patient experience feedback. This dashboard was only accessible to participants in the study; therefore, usability was evaluated on the same participants from the co-design process, inviting reporting bias on the final SUS score. Another important limitation that has implications beyond this particular co-design study is the potential for the idealization of the work context by the staff involved. When staff are taken off the ward and given some time and space to be involved in co-designing an intervention, which later they will deliver in a busy ward, they are not necessarily able to anticipate the difficulties that they will face. Alternatively, they may ignore these challenges because they are fearful of admitting to them in a group setting, particularly in a group that includes patients.

It could be said that we are currently at a key pivotal moment in terms of the patient experience debate in relation to both national and local policy and what is occurring *on the ground*. This is because there is an ever-clearer and acknowledged push for improvement to arise from patient feedback, but individuals

and systems are constrained from doing so. This study has attempted to address the point in the National Institute for Health Research report [14] that there is still uncertainty as to how to present patient experience data in a meaningful and granular way that stimulates local action. An important result and advantage of our study's approach is that it draws together very large FFT data into a thematically driven, simple visual display without loss of the nuances that other manually based methods can have, and it can still allow for exploration of the original free-text comments.

Conclusions

The use of visualization techniques such as dashboards is increasing in response to staff needs for summarized, easily interpreted patient information at the point of care. In this study, through a participatory co-design process and usability heuristic evaluation, we developed and refined a dashboard displaying patient experience, namely FFT data, for use by staff and key stakeholders in near real time. The contributions of this study establish guidance for optimizing the design of FFT dashboards that key stakeholders, especially frontline staff, find meaningful and, in turn, support patient-centered care. The impact of this work is being measured in an ongoing trial, the results of which will guide future refinement, integration with electronic health care records, and steps toward dissemination.

Acknowledgments

The authors would like to thank the Imperial College Healthcare National Health Service Trust staff for their participation in the study, the Digital Reference Group for their input in the study protocol and strategic support, and the Imperial Clinical Analytics Research and Evaluation (iCARE) for translating the language analysis output into near real-time visualizations. This work was supported by the National Institute for Health Research Imperial Patient Safety Translation Research Centre. Infrastructure support was provided by the National Institute for Health Research Imperial Biomedical Research Centre. The research was enabled by the iCARE environment and used the iCARE team and data resources. The study funders did not play a role in the study design; the collection, analysis, and interpretation of the data; the writing of the report; or the decision to submit the paper for publication. In addition, researchers were independent from funders, and all authors had full access to all of the data included in this study and can take responsibility for the integrity of the data and accuracy of the data analysis.

Conflicts of Interest

None declared.

References

1. Lawton R, O'Hara JK, Sheard L, Armitage G, Cocks K, Buckley H, et al. Can patient involvement improve patient safety? A cluster randomised control trial of the Patient Reporting and Action for a Safe Environment (PRASE) intervention. BMJ Qual Saf 2017 Aug 03;26(8):622-631 [FREE Full text] [doi: 10.1136/bmjqs-2016-005570] [Medline: 28159854]
2. Coulter A, Locock L, Ziebland S, Calabrese J. Collecting data on patient experience is not enough: they must be used to improve care. BMJ 2014 Mar 26;348(mar26 1):g2225. [doi: 10.1136/bmj.g2225] [Medline: 24671966]
3. Gleeson H, Calderon A, Swami V, Deighton J, Wolpert M, Edbrooke-Childs J. Systematic review of approaches to using patient experience data for quality improvement in healthcare settings. BMJ Open 2016 Aug 16;6(8):e011907 [FREE Full text] [doi: 10.1136/bmjopen-2016-011907] [Medline: 27531733]
4. Sheard L, Peacock R, Marsh C, Lawton R. What's the problem with patient experience feedback? A macro and micro understanding, based on findings from a three-site UK qualitative study. Health Expect 2019 Feb 22;22(1):46-53 [FREE Full text] [doi: 10.1111/hex.12829] [Medline: 30244499]
5. Reeves R, West E, Barron D. Facilitated patient experience feedback can improve nursing care: a pilot study for a phase III cluster randomised controlled trial. BMC Health Serv Res 2013 Jul 04;13:259 [FREE Full text] [doi: 10.1186/1472-6963-13-259] [Medline: 23826970]
6. Flott KM, Graham C, Darzi A, Mayer E. Can we use patient-reported feedback to drive change? The challenges of using patient-reported feedback and how they might be addressed. BMJ Qual Saf 2017 Jun 20;26(6):502-507. [doi: 10.1136/bmjqs-2016-005223] [Medline: 27325796]
7. Robert G, Cornwell J. Rethinking policy approaches to measuring and improving patient experience. J Health Serv Res Policy 2013 Apr 01;18(2):67-69. [doi: 10.1177/1355819612473583]
8. Sheard L, Marsh C, O'Hara J, Armitage G, Wright J, Lawton R. Exploring how ward staff engage with the implementation of a patient safety intervention: a UK-based qualitative process evaluation. BMJ Open 2017 Jul 13;7(7):e014558 [FREE Full text] [doi: 10.1136/bmjopen-2016-014558] [Medline: 28710206]
9. NHS England and NHS improvement guidance: using the Friends and Family Test to improve patient experience. National Health Service. URL: https://www.england.nhs.uk/wp-content/uploads/2019/09/using-the-fft-to-improve-patient-experience-guidance-v2.pdf [accessed 2021-12-26]
10. Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: review of the literature. Int J Med Inform 2015 Feb;84(2):87-100. [doi: 10.1016/j.ijmedinf.2014.10.001] [Medline: 25453274]

11. Francis R, Mid Staffordshire NHS Foundation Trust Public Inquiry. Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. London, United Kingdom: Stationery Office; 2013.
12. Keogh B. Review into the quality of care and treatment provided by 14 hospital trusts in England: overview report. National Health Service. URL: https://www.nhs.uk/nhsengland/bruce-keogh-review/documents/outcomes/keogh-review-final-report.pdf [accessed 2021-12-26]
13. A promise to learn – a commitment to act: improving the safety of patients in England. National Advisory Group on the Safety of Patients in England. 2013. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/226703/Berwick_Report.pdf [accessed 2021-12-26]
14. Improving care by using patient feedback. NIHR Dissemination Centre. 2019. URL: https://content.nihr.ac.uk/nihrdc/themedreview-04327-PE/Patient-Feedback-WEB.pdf [accessed 2021-12-26]
15. Docherty SL, Vorderstrasse A, Brandon D, Johnson C. Visualization of multidimensional data in nursing science. West J Nurs Res 2017 Jan;39(1):112-126 [FREE Full text] [doi: 10.1177/0193945916672448] [Medline: 27756850]
16. Boyd H, McKernon S, Mullin B, Old A. Improving healthcare through the use of co-design. N Z Med J 2012 Jun 29;125(1357):76-87. [Medline: 22854362]
17. Why encouraging NHS staff to think differently is good for the nation's health. Design council and department of health and social care. 2008. URL: https://www.england.nhs.uk/wp-content/uploads/2016/05/nursing-framework.pdf [accessed 2022-01-24]
18. Nielsen J, Mack RL. Usability Inspection Methods. Hoboken, New Jersey: Wiley; 1994.
19. Sharp H, Preece J, Rogers Y. Interaction Design: Beyond Human-Computer Interaction. Hoboken, New Jersey: Wiley; 2019.
20. Khanbhai M, Anyadi P, Symons J, Flott K, Darzi A, Mayer E. Applying natural language processing and machine learning techniques to patient experience feedback: a systematic review. BMJ Health Care Inform 2021 Mar;28(1):e100262 [FREE Full text] [doi: 10.1136/bmjhci-2020-100262] [Medline: 33653690]
21. NHS patient experience framework. Department of Health and Social Care. 2012. URL: https://www.gov.uk/government/publications/nhs-patient-experience-framework [accessed 2021-12-26]
22. Boyd H. Health Service Co-design: Working with Patients to Improve Healthcare Services : Guide and Toolkit. Auckland: Waitemata District Health Board; 2010.
23. Dowding D, Merrill J. The development of heuristics for evaluation of dashboard visualizations. Appl Clin Inform 2018 Jul;9(3):511-518 [FREE Full text] [doi: 10.1055/s-0038-1666842] [Medline: 29998455]
24. 10 usability heuristics for user interface design. Nielsen Norman Group. 1994. URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2021-12-26]
25. Tory M, Möller T. Evaluating visualizations: do expert reviews work? IEEE Comput Graph Appl 2005 Sep;25(5):8-11. [doi: 10.1109/mcg.2005.102] [Medline: 16209163]
26. Lewis J, Sauro J. Item benchmarks for the system usability scale. J Usability Stud 2018 May;13(3):158-167. [doi: 10.5555/3294033.3294037]
27. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Human Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
28. Weggelaar-Jansen AM, Broekharst DS, de Bruijne M. Developing a hospital-wide quality and safety dashboard: a qualitative research study. BMJ Qual Saf 2018 Dec;27(12):1000-1007 [FREE Full text] [doi: 10.1136/bmjqs-2018-007784] [Medline: 29950323]
29. Ghazisaeidi M, Safdari R, Torabi M, Mirzaee M, Farzi J, Goodini A. Development of performance dashboards in healthcare sector: key practical issues. Acta Inform Med 2015 Oct;23(5):317-321 [FREE Full text] [doi: 10.5455/aim.2015.23.317-321] [Medline: 26635442]
30. Pauwels K, Ambler T, Clark BH, LaPointe P, Reibstein D, Skiera B, et al. Dashboards as a service: why, what, how, and what research is needed? J Serv Res 2009 Aug 17;12(2):175-189. [doi: 10.1177/1094670509344213]
31. De Vito Dabbs A, Myers B, Mc Curry KR, Dunbar-Jacob J, Hawkins R, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
32. Rivas C, Tkacz D, Antao L, Mentzakis E, Gordon M, Anstee S, et al. Automated analysis of free-text comments and dashboard representations in patient experience surveys: a multimethod co-design study. Health Serv Deliv Res 2019;7(23):1-160 (forthcoming) [FREE Full text] [doi: 10.3310/hsdr07230]
33. Sheard L, Marsh C, Mills T, Peacock R, Partridge R, Gwilt I, et al. Using patient experience data to develop a patient experience toolkit to improve hospital care: a mixed-methods study. Health Serv Deliv Res 2019 Oct;7(36). [doi: 10.3310/hsdr07360]

Abbreviations

FFT: Friends and Family Test
iCARE: Imperial Clinical Analytics Research and Evaluation

NHS: National Health Service
QI: quality improvement
SUS: System Usability Scale

Edited by R Kukafka, G Eysenbach; submitted 11.02.21; peer-reviewed by A Gillespie, F Ghezelbash, F Ghezelbash; comments to author 13.04.21; revised version received 12.06.21; accepted 08.10.21; published 03.02.22.

Please cite as:
Khanbhai M, Symons J, Flott K, Harrison-White S, Spofforth J, Klaber R, Manton D, Darzi A, Mayer E
Enriching the Value of Patient Experience Feedback: Web-Based Dashboard Development Using Co-design and Heuristic Evaluation
JMIR Hum Factors 2022;9(1):e27887
URL: https://humanfactors.jmir.org/2022/1/e27887
doi:10.2196/27887
PMID:35113022

Original Paper

User-Centered Development and Testing of the Online Patient-Reported Outcomes, Burdens, and Experiences (PROBE) Survey and the myPROBE App and Integration With the Canadian Bleeding Disorder Registry: Mixed Methods Study

Federico Germini[1,2], MSc, MD; Victoria Borg Debono[1], BHSc, MSc, PhD; David Page[3], BEd; Victoria Zuk[1], MSc; Alexandra Kucher[4], MSc; Chris Cotoi[1], BEng, EMBA; Nicholas Hobson[1], DIT; Michael Sevestre[5], MCS; Mark W Skinner[1,6], JD; Alfonso Iorio[1,2], MD, PhD; PROBE Investigators[7]

[1]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

[2]Department of Medicine, McMaster University, Hamilton, ON, Canada

[3]Canadian Hemophilia Society, Montreal, QC, Canada

[4]Patient Outcomes Research Group Ltd, Washington, DC, United States

[5]Design2Code Inc, Waterloo, ON, Canada

[6]Institute for Policy Advancement Ltd, Washington, DC, United States

[7]see Acknowledgments, Hamilton, ON, Canada

Corresponding Author:
Federico Germini, MSc, MD
Department of Health Research Methods, Evidence, and Impact
McMaster University
1280 Main Street West, 2C Area
Hamilton, ON, L8S 4K1
Canada
Phone: 1 905 525 9140 ext 26771
Email: germinif@mcmaster.ca

Abstract

Background: The Patient-Reported Outcomes, Burdens, and Experiences (PROBE) questionnaire is a tool for assessing the quality of life and disease burden in people living with hemophilia.

Objective: The objectives of our study were (1) to assess the needs of relevant stakeholders involved in the use of PROBE, (2) to develop the software infrastructure needed to meet these needs, and (3) to test the usability of the final product.

Methods: We conducted a series of semistructured interviews of relevant stakeholders, including PROBE investigators, people with hemophilia, and representatives of the sponsor. Based on these, we developed an online survey and a mobile app for iOS and Android. A user group evaluated the final product using the System Usability Scale (SUS) and an open feedback framework.

Results: The online survey was updated, and the myPROBE app for mobile devices and a new application programming interface were developed. The app was tested and modified according to user feedback over multiple cycles. The final version of the app was released in July 2019. Seventeen users aged 23 to 67 years evaluated the final version of the app using the SUS. The median (first, third quartile) SUS score for the app was 85 (68, 88) out of 100. The newly introduced functionalities were as follows: (1) capability to longitudinally track repeated fillings of the questionnaire at different time points by the same participant (as opposed to anonymous completion); (2) linking of the questionnaire with hemophilia registries, starting with the Canadian Bleeding Disorders Registry as a proof of concept; (3) removing or adding questions as needed; and (4) sending notifications to the users (eg, reminders). A new secure database was built for securely storing personal information separately from the questionnaire data. The PROBE online survey is currently available in 96 countries and 34 languages.

Conclusions: The online survey was updated successfully, and the myPROBE app was developed, with a SUS score of 85 (out of 100). The app has been released in 81 countries and 34 languages. This will facilitate data collection for research and advocacy purposes, and the use of this tool in everyday clinical practice.

(JMIR Hum Factors 2022;9(1):e30797) doi:10.2196/30797

KEYWORDS

health-related quality of life; EQ-5D; mobile app; Patient-Reported Outcomes, Burdens, and Experiences (PROBE); hemophilia; mobile health; mHealth; eHealth; telehealth; user-centered design

Introduction

Background

What is Hemophilia

Hemophilia is an inherited X-linked bleeding disorder. Hemophilia A is characterized by a deficiency in the clotting factor VIII, while Hemophilia B is a deficiency in factor IX. Given the reduced ability to form clots, people living with hemophilia experience an increased frequency and duration of bleeding events, which tend to occur mostly within the joints or muscles [1]. The standard of care for hemophilia treatment involves infusing factor replacement to increase factor concentrations in blood [1]. Another option for prophylaxis in people living with hemophilia A is subcutaneous infusions of emicizumab, a bispecific antibody that mimics the function of factor VIII [2].

Quality of Life in Hemophilia

Limiting the effects of bleeding episodes is important for maintaining patient quality of life, as bleeds cause deterioration of the joints and can result in pain and disability [3]. The quality of life of people living with hemophilia and the burden of the disease are also affected by other aspects, for example, the burden of adhering to a regular prophylactic regimen and comorbidities, particularly bleeding-related arthropathy and transfusion-transmitted infections with hepatitis B virus, hepatitis C virus, or HIV [4]. Outcome assessment in hemophilia often comprises clinical measures, such as number of target joints, number of emergency room visits, and hospital length of stay [5]. In recent years, patient-reported outcomes (PROs) have been increasingly used to measure additional metrics that capture a patient's perspective of their own health [6,7]. The assessment of PROs involves asking an individual to assess a variety of domains, such as pain, disability, function, and satisfaction with treatment. This, when combined with clinical outcomes, provides a more comprehensive understanding of the disease and available treatments [8]. Multiple PROs and quality of life measuring instruments exist in hemophilia, although they have not involved patients throughout the development of the questionnaire, despite clear guidance to do so [9]. The Patient-Reported Outcomes, Burdens, and Experiences (PROBE) questionnaire was developed for patients by patients, with the support of an epidemiologist, an expert in outcome research, an expert in management science, and economics-engineering systems [10]. The PROBE questionnaire has undergone assessment for feasibility, validity [10], and test-retest reliability (including cross-validation of paper and online versions) [11], and evaluation of other psychometric properties [12]. The finalized version of the PROBE questionnaire is composed of 29 questions with additional subquestions depending on responses, which have been divided into 4 sections, focusing on personal demographics, general health problems, disease-specific health problems, and the EuroQol 5 dimensions 5 levels (EQ-5D-5L) and Visual Analog Scale (VAS). Before this project, PROBE had been tested across 21 countries, and was available in 11 languages (with 20 localized versions worldwide). It has been shown that PROBE is a valid questionnaire for the evaluation of PROs in people living with hemophilia and a control population [12]. Its discriminative properties allow its use in clinical trials, longitudinal studies, health technology assessment studies, routine clinical care, or disease registries. Before this project, the questionnaire could be filled on the PROBE website anonymously by people living with hemophilia and controls (individuals without a bleeding disorder) alike. This system did not allow for longitudinal data collection. If the same individual returned at a later date to take the questionnaire again, the new questionnaire was not linked to the previous one. The PROBE investigators assessed that the capacity to collect longitudinal data and to link responses to other databases would provide benefits to physicians, people living with hemophilia, and researchers. By doing so, if agreed by the participant, individual responses could be tracked, and progress could be monitored and evaluated.

Digital Health Interventions for Chronic Conditions

The acceptability and feasibility of digital health interventions in chronic conditions (including rare diseases) have already been proven, while evidence about the efficacy of such interventions on patient-oriented outcomes is not definitive [13-16]. In our case, we aimed at eliciting the measurement of PROs. This is not an intervention directly aimed at affecting outcomes, even though this can be an indirect effect. Longitudinal PRO data could support decisions surrounding novel treatments or treatment schedule changes, provide important insights on the changes in quality of life following certain events (eg, bleeds, surgery, or a change in treatment regimen), or provide a tool for physicians or health systems to track patient outcomes over time. Different solutions have been proposed for the digital collection of PRO measures, including modifying the architecture of the electronic health record to integrate data collection from different mobile apps [17]. However, once implemented, the usability of these systems needs to be tested, as changes might be needed if the results are unsatisfactory [18,19].

The backbone of this project was making prospective longitudinal data collection of the PROBE questionnaire possible, by developing a stand-alone individual longitudinal PROBE modality and integrating the data collection of PROBE with other databases. The Canadian Bleeding Disorders Registry (CBDR) was used to demonstrate the feasibility of registry integration, and the usability of the app was tested.

Objectives

The objectives of the project were as follows: (1) to assess the needs of the relevant involved stakeholders, (2) to develop the software infrastructure needed to meet those needs, and (3) to test the usability of the final product with potential users.

Methods

Study Phases

The study was organized in the following phases: (1) needs assessment, (2) identification of software specifications, software development, and beta testing, and (3) usability test of the final product.

Needs Assessment

To plan for the development of the software infrastructure and design of the PROBE project, we conducted a series of semistructured interviews with a convenience sample of front-end and back-end users, including all relevant stakeholders. In detail, back-end users were the PROBE investigators, representatives of the sponsor, and researchers, and front-end users were people with hemophilia. The interviewees were based in Canada, Ireland, Italy, Switzerland, and the United States of America. In preparation for the interview, we asked the interviewees to complete the PROBE questionnaire on the website, unless they had recently taken it. The semistructured interview guide is provided in Textbox 1. Characteristics of the stakeholders are provided in Multimedia Appendix 1.

Textbox 1. Guide of the semistructured interview for the needs assessment phase.

1. What's your role in Patient-Reported Outcomes, Burdens, and Experiences (PROBE), if any?
2. What do you like of the current PROBE website?
3. What is that you do not like of it?
4. What functionalities do you think are missing in the PROBE website/paper version?
5. What future do you envision for PROBE, especially in terms of its use for advocacy, clinical research, and clinical activity?
6. Do you think updating the website would be enough for these scopes, or a mobile app is needed?

Software Specifications, Development, and Beta Testing

We put together a technical group involving (1) 3 programmers on staff at the Health Information Research Unit at McMaster University, (2) an external consultant (Design2Code [D2C]) based in Waterloo and skilled in mobile app development, and (3) stakeholder representatives. The needs identified in the previous phase were used to guide the creation of 2 alternative plans for software development: updating the website and creating a mobile app or only updating the website to meet those needs. The 2 plans were discussed within the technical group using the Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis process to compare the different alternatives, including their costs and the time needed for development. The product of this phase was a justification of the choice and a detailed description of the technical specifications for realizing the PROBE suite. The main areas of work were defining if an app was needed and which characteristics were required, and understanding how the existing PROBE website, database, and application programming interface (API) needed to be modified.

All members of the technical team and selected PROBE investigators were invited to test the software under development and provide feedback. Testing included 3 sessions of group testing and independent individual testing. The first round of testing and feedback was based on a mock-up of the product. The second round was based on the beta version of the new website and the app, released on TestFlight for iOS and on the testing environment of Google Play. These products were working with a fake development environment (API and database). The third round was based on the first release of the website and mobile app, using a test modality on the production database.

Usability Testing

Once the software was stable in its first version, which took 3 releases, formal usability testing was performed. A user group composed of a convenience sample of patients, representatives of the sponsor from the same countries specified above, and graduate students from McMaster University evaluated the final product with the System Usability Scale (SUS) [20] and an open feedback framework. The SUS is composed of 10 questions (reported in Multimedia Appendix 2) to be answered on a 5-point Likert scale, from "strongly agree" to "strongly disagree."

Statistical Analysis

The SUS score can range from 0 (worst usability possible) to 100 (best usability possible), calculated as suggested by John Brooke as follows: "to calculate the SUS score, first sum the score contributions from each item. Each item's score contribution will range from 0 to 4. For items 1, 3, 5, 7, and 9, the score contribution is the scale position minus 1. For items 2, 4, 6, 8, and 10, the contribution is 5 minus the scale position. Multiply the sum of the scores by 2.5 to obtain the overall value of the SUS" [20]. The quantitative results of phase 3 of the project were presented using measures of central tendency and dispersion, or counts (frequencies) as appropriate. Data were analyzed using STATA/SE V.16.0 (StataCorp).

Ethical Approval

The PROBE project was approved by the Hamilton Integrated Research Ethics Board (HIREB; application number 7492).

Results

Overview

The project started in November 2017. The needs assessment phase lasted 5 months. The project development phase started in March 2018. The usability test started in March 2019. The myPROBE app was officially released in July 2019.

Needs Assessment

As a result of the semistructured interviews, a complete list of required functionalities was compiled. These functionalities are reported with the respective explanations in Textbox 2. The main needs that were identified included the longitudinal repetition of the questionnaire, linkage with other databases (starting with the CBDR as a proof of concept), and flexibility to allow adding or removing questions. There was a general agreement that these key functionalities were essential for the uptake of PROBE in clinical activity and clinical studies. Moreover, from the perspective of users, the possibility to complete the questionnaire on portable devices (smartphones and tablets) and to save an incomplete questionnaire and complete it later were identified as important features.

Textbox 2. List of functionalities for the new Patient-Reported Outcomes, Burdens, and Experiences (PROBE) suite.

Longitudinal repetition of the questionnaire

Being able to identify two or more sets of answers to the Patient-Reported Outcomes, Burdens, and Experiences (PROBE) questionnaire as coming from the same user. This is key for the use of the questionnaire in clinical activities, in clinical studies, and to assess the responsiveness of the PROBE tool.

How the goal was achieved: Users are now asked to create a username and a password and to login to the system before completing the questionnaire.

Linking PROBE data with other databases

Many registries on patients with bleeding disorders are available around the world. Linking these databases with PROBE data would allow using PROBE in clinical activity and enhance its use in research. Moreover, an increasing number of clinical trials are using PROBE to assess the efficacy of treatments.

The Canadian registry (Canadian Bleeding Disorders Registry [CBDR]) was selected to demonstrate proof of concept, leaving open the possibility to later add other registries and study databases in the future.

How the goal was achieved: Single sign-on with OAuth 2.0 technologies.

Turning modules on and off

When specific information is already available in a linked database or form previous questionnaires, it is not efficient to ask again, so some questions could be removed. If additional information must be collected for a specific study, some questions could be added.

How the goal was achieved: A survey manager allows the creation of new sections and questions for the PROBE questionnaire. A template builder allows to group sections and questions to generate different questionnaire templates.

Completing the questionnaire on a portable device

Smartphones and tablets are being more commonly used and are preferred to laptops and personal computers from many users. Moreover, smartphones are the only available devices to access the internet for the majority of users in low-income countries.

How the goal was achieved: A mobile app for iOS and Android was created.

Saving an incomplete questionnaire and completing it later

To minimize the loss of data and enhance the user experience.

How the goal was achieved: Data are saved locally (and submitted to the database, if a connection is available) every time a user answers a question.

Send notifications to a patient

For example, when it is time to repeat the questionnaire (eg, after 1 year, or after a bleed or an invasive procedure has been recorded in a linked database).

How the goal was achieved: For now, only email notifications can be sent to the users.

Ensuring continuity of data collection

The data from the PROBE questionnaire need to keep flowing to the existing PROBE database.

How the goal was achieved: Anonymized data are stored in the PROBE database, and personal identifiers are stored separately.

Recording the time spent completing the questionnaire

This was, for the PROBE investigators, an important measure of the questionnaire's feasibility.

How the goal was achieved: The time elapsed from the questionnaire loading to its submission is recorded, and the information is stored in the database.

Recording the questionnaire completion rate

Again, to assess the feasibility of the questionnaire, the PROBE investigators need to track the number of users starting the questionnaire and the number of users submitting it in general and among users asked to complete the questionnaire via notifications.

How the goal was achieved: Every time a questionnaire is started, the answers are stored in the database. A variable identifies the questionnaires that have been submitted.

Software Specifications and Development

App Versus Device-Responsive Website

Based on the results of the needs assessment and following the structured approach described as module 2 of the study, the PROBE team met with D2C to decide if an app was needed or if a device-responsive website was sufficient to achieve the objectives. The SWAT analysis results between the 3 plans were translated in pros and cons of having an app on top of the website, which have been summarized in Table 1. The team determined that an app was needed, primarily to facilitate the accessibility of the content, to send notifications to users when

it is time to repeat the questionnaire (eg, after 6 months or when a bleed occurs), and to allow, in the future, leveraging of smartphone features like physical activity tracking or access to the camera. Furthermore, off-line questionnaire completion would not be possible on a website, while this was perceived as important, especially from the perspective of users from less developed countries where internet access or bandwidth is limited. Therefore, to meet these needs, it was agreed that an app for Android and iOS environments was required.

Table 1. Pros and cons of a mobile app and a device-responsive website.

Feature	App pro/con	Website pro/con
Accessibility from a portable device	Easily accessible through an icon.	Requires the user to save the bookmark on the home screen.
Notifications (eg, to repeat the questionnaire)	Available also while not using the app.	Only available while using the website or though emails/text messages.
Offline questionnaire completion	Connection required only to download and submit the questionnaire. Completion can happen offline.	Not accessible offline.
Leveraging smartphone features (eg, step count or camera)	Easier to achieve.	Harder to achieve.
General accessibility	Only accessible through a portable device. Specific operating system versions needed (eg, iOS and Android).	Accessible through a browser from a computer and a portable device.
Costs for development, maintenance, and update	In our case, a website was needed, so costs for the app would be added on top of website expenses.	Not having a website was not an option for the investigators.

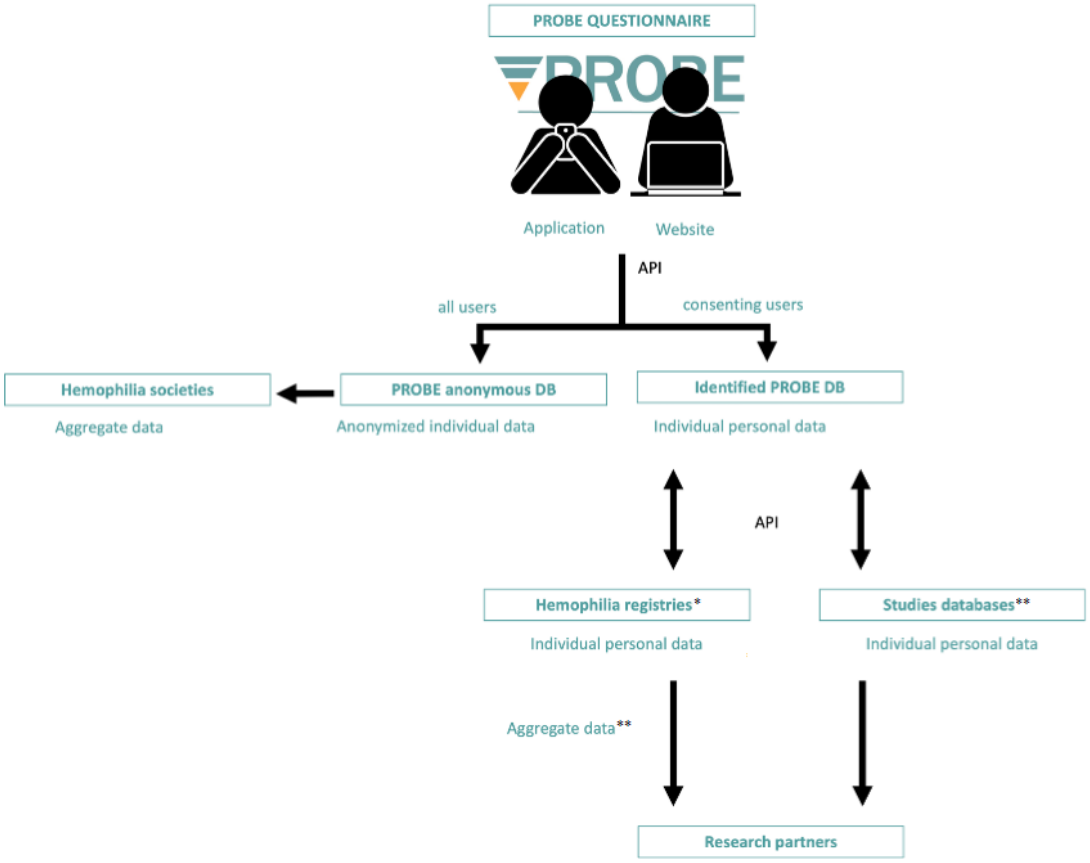
Log-in Flow

To allow for the longitudinal repetition of the PROBE questionnaire, log-in options were required. Retaining the anonymous completion option was deemed important, as was adding an email-based login and a single sign-on (SSO) option for people living with hemophilia to participate through their national registry patient app (eg, myCBDR). The CBDR option was only made available to users selecting “Canada” as their home country.

Database Structure

The data flow from completion to data analysis and study reports is described in Figure 1. Individual data needed to flow from the app and the website to the PROBE database. For security reasons, personal information (email address or CBDR identifier) had to be stored separately from the PROBE questionnaire’s data. The capability to allow part of the data to flow bidirectionally from the PROBE database to others, like hemophilia registries, was needed.

Figure 1. PROBE data flow. The pictures "using smartphone" and "using laptop" are by Llisole from the Noun Project (https://thenounproject.com). API: application programming interface; DB: database; PROBE: Patient Reported Outcomes, Burdens and Experiences. *Starting with the Canadian Bleeding Disorders Registry, with other registries in the future, **Based on future agreements.



API

Before this project, the PROBE website was communicating directly with the PROBE database. To implement the new functionalities, an API was required. Having an API in place allows (1) sending data to the databases from both the app and the website, (2) authenticating users with a dedicated email and password or with a token (eg, obtained using CBDR credentials or coming back to the PROBE website after the first log-in), and (3) supporting different versions of the questionnaire and other data (eg, notifications or calculation and report of PROBE and EQ-5D scores) from the database to the app and website. The technical specifications for the PROBE suite are reported in Multimedia Appendix 3 and Multimedia Appendix 4.

In collaboration with the PROBE investigators, the team at McMaster University and D2C developed an online survey using Microsoft.Net technologies and an app for iOS and Android using react-native. For the duration of the development phase, the McMaster University team and D2C met monthly to discuss progress and to find solutions to unanticipated problems. The PROBE investigators and the sponsor were involved as needed. Textbox 2 reports the solutions implemented to realize the main system functionalities. A sign-in and log-in interface was created. The SSO with MyCBDR credentials was implemented using OAuth 2.0 technologies. A survey manager and a template builder were implemented to allow creating new versions of the PROBE questionnaire by adding or removing sections and questions.

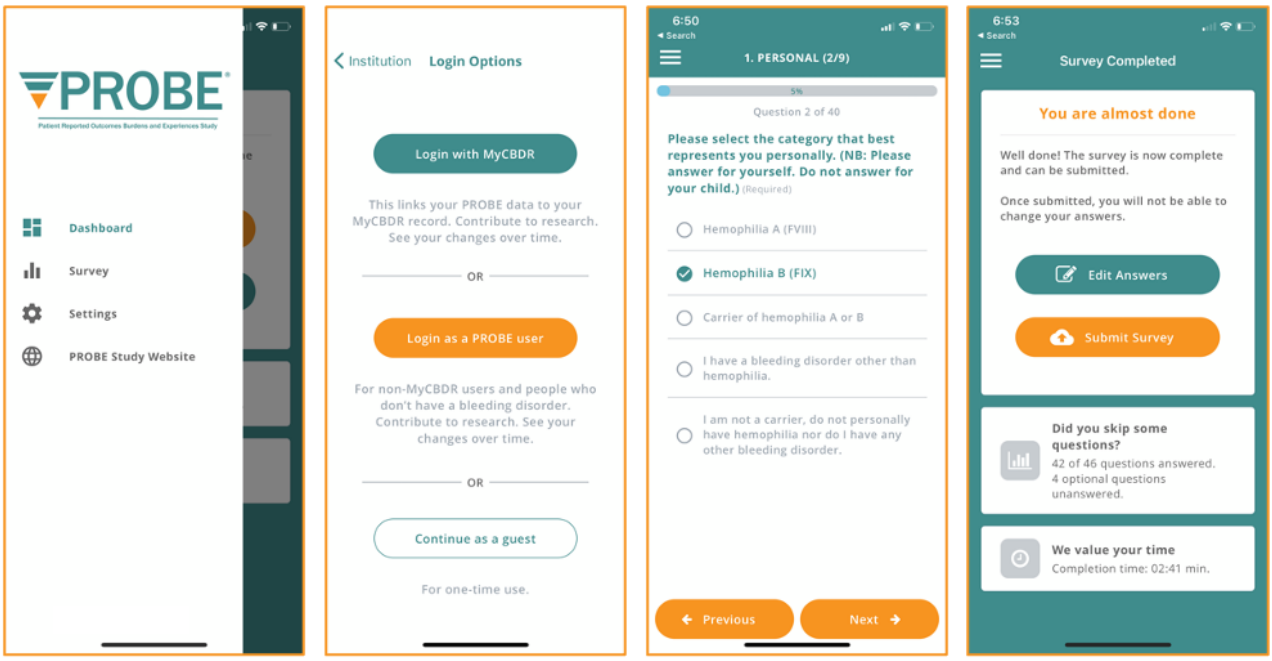
The 3 cycles of testing of the app and website allowed bugs and required fixes to be identified. Feedback from beta-testers led to improvements in the user experience, for example, adding autoscroll for pages displaying more than one question, changing the app buttons and graphics when they were perceived as not being clear, and adding descriptive text (eg, feet, inches, and pounds to the height and weight question for countries not using the metric system).

Usability Testing

Once a stable version was achieved, 17 users aged 23 to 67 years evaluated the app using the SUS. The median (first, third quartile) SUS score for the app was 85 (68, 88) out of 100. Based on the open-ended comments, most users indicated satisfaction. The only major edit requested was to allow data validation when possible. The app was modified accordingly (eg, "age" was restricted to be a number between 0 and 100).

Figure 2 shows screenshots from the first release of the myPROBE app.

Figure 2. Screenshots of the myPROBE app.



Discussion

Principal Findings

With a user-centered approach, careful needs assessment, and testing, a user-friendly mobile app was developed that allows longitudinal completion of the PROBE questionnaire, administration of different questionnaires created ad hoc, and linkage with other databases through SSO. The myPROBE app was released on the Apple Store and Google Play Store in 81 countries and 34 languages. Testers favorably rated the usability of the app with a median score of 85 (out of 100).

Strengths and Limitations

Time and resources for this project were limited. Therefore, some functionalities that users requested could not be implemented. We were successful in implementing SSO through MyCBDR, and recently added SSO with the Mexican Registry of Bleeding Disorders. However, for now, we are not supporting SSO with credentials from other services, like Apple or Google. Implementing this option might have relieved users from having to remember a new username and password specifically for PROBE, further increasing the usability of the app. Caregivers and researchers have a growing interest in linking quality of life (QoL) data with data on physical activity in people living with hemophilia. Having a mobile app paves the way for passive data collection through a wearable device (like a smartwatch). However, this is not yet possible for the myPROBE app. Some functionalities offered on the myPROBE website are not offered on the myPROBE app. In particular, the website allows users to download and share their questionnaire results and offers better support for participation in studies, with the possibility of administering different questionnaires to specific users. Additionally, functionalities have been recently developed but have not, as yet, been implemented on the myPROBE app. Retrocompatibility was limited to Android 5+ and iOS 10+. This might restrict access to the app, especially in low-income countries where older mobile devices may be commonly used. The strengths of our study include the multidisciplinary nature of the team, which involved experts in information technology and health research methodology, hemophilia treaters, people working in the industry, and, perhaps most importantly, people with hemophilia. We believe that the user-centered approach with early involvement of final users in the development of our product was key to determine the good usability of the app. The widespread distribution of the app and its translation in more than 30 languages will favor its uptake and will foster new feedback on how to further improve it.

Comparison With Prior Work

To the best of our knowledge, myPROBE is the only mobile app for completing a QoL tool currently available on the market for people living with hemophilia.

The SUS has been widely up-taken by academics (6932 citations in Google Scholar on February 12, 2019) and practitioners, and it has been used in a large variety of settings, ranging from safety signs [21] to websites [22]. The SUS showed good to excellent psychometric properties in different settings [23]. Comparing the scores with benchmark measures [23], the myPROBE app was rated A+ (96th to 100th percentile).

In terms of mobile apps for collecting PRO data in patients with chronic conditions other than hemophilia, previous studies reported conflicting results. Welbie et al tested the usability of an app for PROs in physical therapy patients and found that users where overall satisfied with the usability of the app, but

the app required some changes in navigation through questions and how to insert and edit answers [19]. The authors acknowledged that they would again test the usability of the app after implementing such changes. We found similar issues during our testing phase, and our approach was to implement such changes before testing usability. This confirms how it is critical to include direct patients' inputs in the development of digital health interventions [13,24], and how key considerations for end users should be sought early on in the process of app or digital health intervention design to ensure short- and long-term engagement [25]. For example, it was shown how aspects that might be considered less important by researchers, like design, communication style, and user ratings, can be important for engagement [26]. Steele Gray et al investigated the usability of an app to report PROs in complex chronic disease and disability [18]. They highlighted issues with usability related to the frequency of questionnaire administration and actual use of the data in a clinical setting. The former issue is not directly related to the usability of a mobile app but more to the PRO tool in general. The former is an aspect that we did not explore in this study, as we wanted to focus on the usability of the app to collect data before moving to the use of the data in a clinical setting. Both the studies used qualitative approaches to explore the usability of the apps. On one side, this is valuable and allows more in-depth feedback and exploration of aspects that are not measurable. On the other hand, using a quantitative tool allows quantifying the usability and performing comparisons with other tools. For these reasons, we opted for a mixed methods approach, with a qualitative approach for development and initial feedback, and a quantitative approach for obtaining a measure of usability.

Future Directions

We are continually expanding the global reach of PROBE with new releases of the myPROBE app, including new translations, thus making the app available in more countries and affording the opportunity to support the integration with other bleeding disorder registries. A frequently asked questions (FAQs) handout was produced in multiple languages, describing the scope of the PROBE project and app functionalities (Multimedia Appendix 5).

An SSO with the World Federation of Hemophilia World Bleeding Disorders Registry is being developed as we write this paper. A study to prove the test-retest reliability of PROBE when administered through the mobile app and the website is in process. Other future directions for the app will be addressing the above-mentioned limitations. In particular, funding for supporting passive data collection for physical activity is being pursued. There is growing interest in how to use PROBE data in clinical activity. We are working on offering the possibility to provide caregivers of people living with hemophilia (when people living with hemophilia consent) access to PROBE data, and to compare individual data with group data from other users [27]. For example, people living with hemophilia or their caregivers might want to know how they compare to people in the same age range and country, with or without hemophilia. Use of the PROBE questionnaire and benchmark data sets might prompt reflections on current problems and goals of care, potentially improving the care of people living with hemophilia. The integration with bleeding disorder registries offers the possibility of shortening the questionnaire and avoiding asking for information already available (eg, year of birth). Moreover, it would be possible to use data from these registries to prompt event-based completion of the questionnaire, for example, after a bleeding event is registered by a user or after a significant change in treatment access or standard of care in a country. Notifications might also be used to remind the user when it is time to repeat the questionnaire (eg, every 6 months).

In general, we believe that future studies on digital health interventions should involve end users in the early development phases, to ensure good usability and engagement. Moreover, adopting a widespread tool to measure usability would allow comparisons between different digital health interventions. Once developed, the real-world use of these tools should be assessed in terms of usability in a clinic setting ideally to see how these interactions can translate into patient care changes and if this can also affect outcomes. Lastly, economic evaluations should be performed to support the use of digital health interventions, and this aspect has been rarely investigated to date [28,29].

Conclusions

The PROBE online survey was updated successfully, and the myPROBE app was developed using a user-centered approach. This allows digital administration of the PROBE questionnaire and other questionnaires, and adoption of SSO for ease of use and linkage to other databases. In the first months after the product's release in 81 countries and 34 languages, the responses from testers and users have been largely positive. The median SUS score (85/100) compares well with previously published benchmark measures. We believe that this is a crucial step toward facilitating the use of this PRO tool in research and everyday patient care. This will contribute to pursuing the objectives of the PROBE project, including building a robust evidence base for comparative effectiveness, outcome research, evidence-based decision making, and advocacy.

Acknowledgments

McMaster University received funding from Hoffmann-La Roche Limited for the realization of the myPROBE app. The Patient-Reported Outcomes, Burdens, and Experiences (PROBE) study received research support from Bayer, CSL, Novo Nordisk, Roche, Sanofi, Sobi, and Takeda. The PROBE investigators were Mark W Skinner, Chatree Chai-Adisaksopha, Randall G Curtis, Neil Frick, Alfonso Iorio, Michael B Nichol, Declan Noone, Brian O'Mahony, David Page, and Jeffrey S Stonebraker.

Conflicts of Interest

FG and AI's institution (McMaster University) received research funds from NovoNordisk, Roche, Takeda, Bayer, Pfizer, BioMarin, CSL, Freeline, Grifols, Octapharma, Sanofi, Spark, and Uniqure. MS received consultancy honoraria and research funding from Bayer, Biomarin, and Genetech, honoraria and research funding from Novo Nordisk, Roche, Sanofi, and Takeda, and research funding from Freeline Therapeutics, Sobi, uniQure, and other from Pfizer and Spark Therapeutics.

Multimedia Appendix 1
Characteristics of the interviewees for the needs assessment phase (semistructured interviews).
[DOCX File , 24 KB - humanfactors_v9i1e30797_app1.docx]

Multimedia Appendix 2
System usability scale.
[PDF File (Adobe PDF File), 254 KB - humanfactors_v9i1e30797_app2.pdf]

Multimedia Appendix 3
Technical requirements specification for the Patient-Reported Outcomes, Burdens, and Experiences (PROBE) service (excluding the PROBE app).
[PDF File (Adobe PDF File), 224 KB - humanfactors_v9i1e30797_app3.pdf]

Multimedia Appendix 4
Technical requirements specification for the Patient-Reported Outcomes, Burdens, and Experiences (PROBE) mobile app for iOS and Android.
[PDF File (Adobe PDF File), 214 KB - humanfactors_v9i1e30797_app4.pdf]

Multimedia Appendix 5
myPROBE app handout.
[PDF File (Adobe PDF File), 303 KB - humanfactors_v9i1e30797_app5.pdf]

References

1. Srivastava A, Santagostino E, Dougall A, Kitchen S, Sutherland M, Pipe S, WFH Guidelines for the Management of Hemophilia panelistsco-authors. WFH Guidelines for the Management of Hemophilia, 3rd edition. Haemophilia 2020 Aug;26 Suppl 6:1-158. [doi: 10.1111/hae.14046] [Medline: 32744769]
2. Mahlangu J, Oldenburg J, Paz-Priel I, Negrier C, Niggli M, Mancuso M, et al. Emicizumab Prophylaxis in Patients Who Have Hemophilia A without Inhibitors. N Engl J Med 2018 Aug 30;379(9):811-822. [doi: 10.1056/NEJMoa1803550] [Medline: 30157389]
3. Rodriguez-Merchan EC. Musculoskeletal complications of hemophilia. HSS J 2010 Feb;6(1):37-42 [FREE Full text] [doi: 10.1007/s11420-009-9140-9] [Medline: 19921342]
4. van Balen EC, O'Mahony B, Cnossen MH, Dolan G, Blanchette VS, Fischer K, et al. Patient-relevant health outcomes for hemophilia care: Development of an international standard outcomes set. Res Pract Thromb Haemost 2021 May 06;5(4):e12488 [FREE Full text] [doi: 10.1002/rth2.12488] [Medline: 34027286]
5. Aquino CC, Borg Debono V, Germini F, Pete D, Kempton CL, Young G, et al. Outcomes for studies assessing the efficacy of hemostatic therapies in persons with congenital bleeding disorders. Haemophilia 2021 Mar 06;27(2):211-220. [doi: 10.1111/hae.14247] [Medline: 33550614]
6. Hoos A, Anderson J, Boutin M, Dewulf L, Geissler J, Johnston G, et al. Partnering With Patients in the Development and Lifecycle of Medicines: A Call for Action. Ther Innov Regul Sci 2015 Nov;49(6):929-939 [FREE Full text] [doi: 10.1177/2168479015580384] [Medline: 26539338]
7. The Voice of the Patient: A series of reports from the U.S. Food and Drug Administration's (FDA's) Patient-Focused Drug Development Initiative. FDA. 2016. URL: https://www.fda.gov/media/99237/download [accessed 2022-01-23]
8. Skinner M. WFH: closing the global gap--achieving optimal care. Haemophilia 2012 Jul;18 Suppl 4:1-12 [FREE Full text] [doi: 10.1111/j.1365-2516.2012.02822.x] [Medline: 22726075]
9. Lowe M, Blaser D, Cone L, Arcona S, Ko J, Sasane R, et al. Increasing Patient Involvement in Drug Development. Value Health 2016;19(6):869-878 [FREE Full text] [doi: 10.1016/j.jval.2016.04.009] [Medline: 27712716]
10. Skinner MW, Chai-Adisaksopha C, Curtis R, Frick N, Nichol M, Noone D, et al. The Patient Reported Outcomes, Burdens and Experiences (PROBE) Project: development and evaluation of a questionnaire assessing patient reported outcomes in people with haemophilia. Pilot Feasibility Stud 2018;4:58 [FREE Full text] [doi: 10.1186/s40814-018-0253-0] [Medline: 29497561]

11. Chai-Adisaksopha C, Skinner M, Curtis R, Frick N, Nichol M, Noone D, et al. Test-retest properties of the Patient Reported Outcomes, Burdens and Experiences (PROBE) questionnaire and its constituent domains. Haemophilia 2019 Jan;25(1):75-83 [FREE Full text] [doi: 10.1111/hae.13649] [Medline: 30520534]
12. Chai-Adisaksopha C, Skinner MW, Curtis R, Frick N, Nichol MB, Noone D, et al. Psychometric properties of the Patient Reported Outcomes, Burdens and Experiences (PROBE) questionnaire. BMJ Open 2018 Aug 08;8(8):e021900 [FREE Full text] [doi: 10.1136/bmjopen-2018-021900] [Medline: 30093520]
13. Badawy SM, Kuhns LM. Texting and Mobile Phone App Interventions for Improving Adherence to Preventive Behavior in Adolescents: A Systematic Review. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e50 [FREE Full text] [doi: 10.2196/mhealth.6837] [Medline: 28428157]
14. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text Messaging and Mobile Phone Apps as Interventions to Improve Adherence in Adolescents With Chronic Health Conditions: A Systematic Review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]
15. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: A systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan 19;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]
16. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-Centered eHealth Interventions for Children, Adolescents, and Adults With Sickle Cell Disease: Systematic Review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]
17. Ploner N, Prokosch HU. Integrating a Secure and Generic Mobile App for Patient Reported Outcome Acquisition into an EHR Infrastructure Based on FHIR Resources. Stud Health Technol Inform 2020 Jun 16;270:991-995. [doi: 10.3233/SHTI200310] [Medline: 32570530]
18. Steele Gray C, Gill A, Khan AI, Hans PK, Kuluski K, Cott C. The Electronic Patient Reported Outcome Tool: Testing Usability and Feasibility of a Mobile App and Portal to Support Care for Patients With Complex Chronic Disease and Disability in Primary Care Settings. JMIR Mhealth Uhealth 2016 Jun 02;4(2):e58 [FREE Full text] [doi: 10.2196/mhealth.5331] [Medline: 27256035]
19. Welbie M, Wittink H, Westerman MJ, Topper I, Snoei J, Devillé WLJM. A Mobile Patient-Reported Outcome Measure App With Talking Touchscreen: Usability Assessment. JMIR Form Res 2019 Sep 27;3(3):e11617 [FREE Full text] [doi: 10.2196/11617] [Medline: 31573909]
20. Brooke J. SUS - A quick and dirty usability scale. TBIStaffTraining. URL: http://www.tbistafftraining.info/smartphones/documents/b5_during_the_trial_usability_scale_v1_09aug11.pdf [accessed 2022-01-23]
21. Ng A, Lo H, Chan A. Measuring the usability of safety signs: A use of System Usability Scale (SUS). 2011 Presented at: International MultiConference of Engineers and Computer Scientists 2011, IMECS 2011; March 16-18, 2011; Hong Kong p. 1296-1301.
22. Phillips C, Chaparro B. Visual Appeal vs. Usability: Which One Influences User Perceptions of a Website More? Research in User Experience. 2009. URL: https://researchinuserexperience.wordpress.com/2009/10/15/visual-appeal-vs-usability-which-one-influences-user-perceptions-of-a-website-more/ [accessed 2022-01-23]
23. Lewis JR, Sauro J. Item Benchmarks for the System Usability Scale. J Usability Stud 2018;13(3):158-167 [FREE Full text]
24. Badawy SM, Thompson AA, Liem RI. Technology Access and Smartphone App Preferences for Medication Adherence in Adolescents and Young Adults With Sickle Cell Disease. Pediatr Blood Cancer 2016 May 04;63(5):848-852. [doi: 10.1002/pbc.25905] [Medline: 26844685]
25. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
26. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
27. Jackson S, Sheridan B, Gue D, Lyons S, Waines R, Ireland T, et al. Feasibility of Integrating PROBE (Patient Reported Outcomes, Burdens and Experiences) into a Clinical Care Pathway to Facilitate Transformation to a Value-based Health Care Model in Hemophilia. ISTH Congress. URL: https://tinyurl.com/s2wtmv6n [accessed 2022-01-23]
28. Badawy SM, Kuhns LM. Economic Evaluation of Text-Messaging and Smartphone-Based Interventions to Improve Medication Adherence in Adolescents with Chronic Health Conditions: A Systematic Review. JMIR Mhealth Uhealth 2016 Oct 25;4(4):e121 [FREE Full text] [doi: 10.2196/mhealth.6425] [Medline: 27780795]
29. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017 Feb 2;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]

Abbreviations

API: application programming interface

CBDR: Canadian Bleeding Disorders Registry
D2C: Design2Code
PRO: patient-reported outcome
PROBE: Patient-Reported Outcomes, Burdens, and Experiences
SSO: single sign-on
SUS: System Usability Scale

Edited by A Kushniruk; submitted 28.05.21; peer-reviewed by S Badawy, P Knapp, H Mehdizadeh; comments to author 31.08.21; revised version received 25.10.21; accepted 11.11.21; published 02.03.22.

Please cite as:
Germini F, Borg Debono V, Page D, Zuk V, Kucher A, Cotoi C, Hobson N, Sevestre M, Skinner MW, Iorio A, PROBE Investigators User-Centered Development and Testing of the Online Patient-Reported Outcomes, Burdens, and Experiences (PROBE) Survey and the myPROBE App and Integration With the Canadian Bleeding Disorder Registry: Mixed Methods Study
JMIR Hum Factors 2022;9(1):e30797
URL: https://humanfactors.jmir.org/2022/1/e30797
doi:10.2196/30797
PMID:35234648

Original Paper

Concept Libraries for Repeatable and Reusable Research: Qualitative Study Exploring the Needs of Users

Zahra Almowil[1], MSc; Shang-Ming Zhou[2], PhD; Sinead Brophy[1], PhD; Jodie Croxall[1], PhD

[1]Data Science Building, Medical School, Swansea University, Swansea, Wales, United Kingdom

[2]Centre For Health Technology, Faculty of Health, University of Plymouth, Plymouth, United Kingdom

Corresponding Author:
Zahra Almowil, MSc
Data Science Building
Medical School, Swansea University
Sketty
Swansea, Wales, SA2 8PP
United Kingdom
Phone: 44 07552894384
Email: 934467@swansea.ac.uk

Abstract

Background: Big data research in the field of health sciences is hindered by a lack of agreement on how to identify and define different conditions and their medications. This means that researchers and health professionals often have different phenotype definitions for the same condition. This lack of agreement makes it difficult to compare different study findings and hinders the ability to conduct repeatable and reusable research.

Objective: This study aims to examine the requirements of various users, such as researchers, clinicians, machine learning experts, and managers, in the development of a data portal for phenotypes (a concept library).

Methods: This was a qualitative study using interviews and focus group discussion. One-to-one interviews were conducted with researchers, clinicians, machine learning experts, and senior research managers in health data science (N=6) to explore their specific needs in the development of a concept library. In addition, a focus group discussion with researchers (N=14) working with the Secured Anonymized Information Linkage databank, a national eHealth data linkage infrastructure, was held to perform a SWOT (strengths, weaknesses, opportunities, and threats) analysis for the phenotyping system and the proposed concept library. The interviews and focus group discussion were transcribed verbatim, and 2 thematic analyses were performed.

Results: Most of the participants thought that the prototype concept library would be a very helpful resource for conducting repeatable research, but they specified that many requirements are needed before its development. Although all the participants stated that they were aware of some existing concept libraries, most of them expressed negative perceptions about them. The participants mentioned several facilitators that would stimulate them to share their work and reuse the work of others, and they pointed out several barriers that could inhibit them from sharing their work and reusing the work of others. The participants suggested some developments that they would like to see to improve reproducible research output using routine data.

Conclusions: The study indicated that most interviewees valued a concept library for phenotypes. However, only half of the participants felt that they would contribute by providing definitions for the concept library, and they reported many barriers regarding sharing their work on a publicly accessible platform. Analysis of interviews and the focus group discussion revealed that different stakeholders have different requirements, facilitators, barriers, and concerns about a prototype concept library.

(JMIR Hum Factors 2022;9(1):e31021) doi:10.2196/31021

KEYWORDS

electronic health records; record linkage; reproducible research; clinical codes; concept libraries

Introduction

Background

Health care systems are becoming more digitally focused rather than paper-based and are moving to the use of electronic health records (EHRs) [1]. This means there is a large amount of electronic patient data that can be moved and linked together into safe data repositories to enable researchers and data analysts to query and examine these data effectively [2-5]. The growing availability of electronic patient data offers health care practitioners increased opportunities for secondary use of EHR data to improve the quality of care and research [6-8]. However, the present literature does not describe the barriers that make the use of data and deidentification processes difficult nor does it focus on users' practical needs for data linking [9]. A study observed that "One of the fundamental steps in utilizing this EHRs data is identifying patients with certain characteristics of interest (either exposures or outcomes) via a process known as electronic phenotyping" [10]. Phenotyping is the process of extracting phenotypes from clinical data using computer-executable algorithms [11], and phenotypes are "the measurable biological, behavioural and clinical markers of a condition or disease" [12]. Phenotypes might be as simple as patients with type 2 diabetes or as complex as patients with stage II prostate cancer with urinary urgency but no indications of urinary tract infection [10].

There has been an annual rise at a rate of approximately 20% in primary care research using EHRs in the United Kingdom, which gathers data on general practice from the following databases [13]: Clinical Practice Research Data Link [14], The Health Improvement Network [15], QResearch [16], and Secured Anonymized Information Linkage (SAIL) [17]. However, with different data sets (eg, hospital, general practice, or emergency care), defining a condition is still very subjective, as there are many phenotyping algorithms for identifying the same condition (eg, there are currently 66 ways of defining asthma using routine health data) [18], and interpretation or manipulation of data often requires knowledge of complex programing languages, such as SQL [4]. This means that EHRs are still not accessible to many as their use requires specialized programing skills.

One of the most important factors for reproducible research is the availability of clinical codes in EHR-based research because researchers, clinicians, and health informatics professionals often use them to identify the target population and their specific conditions, known as phenotyping [8,19]. If researchers do not publish the code lists they used (eg, how they were established and the accurate phenotype definitions along with the original research using them), then an essential component of these studies is missing. In the absence of clinical code lists, data analysts would be unable to identify patients with or without conditions [19], and researchers would not be able to compare studies effectively. Even though code lists are available in some studies, researchers often encounter difficulties in retrieving relevant data from code lists created for another research project. Moreover, in specific uncommon conditions, minor errors in the selection of code lists may lead to misclassification of large numbers of patients, leading to biased results [20]. Although using previously developed phenotyping algorithms is often of interest to researchers in many studies, there are many challenges associated with reusing and replicating them effectively [21]. Therefore, it is extremely difficult to assess the validity and transparency of EHR-driven studies [22].

Although researchers request better transparency in sharing clinical code lists [23,24], they face difficulties in obtaining comprehensive code lists from EHR-based research. Although, there are currently no obligations from journals and funding parties to publish code lists, the Strengthening the Reporting of Observational Studies in Epidemiology and Reporting of Studies Conducted Using Observational Routinely Collected Health Data initiatives encourage transparency and open access to publicly available EHR-based research [25-27]. To address these challenges, different data linkage centers in the United Kingdom and other countries, such as Canada, have developed data portals for phenotypes (concept libraries), such as ClinicalCodes.org [22], Clinical Disease Research Using Linked Bespoke Studies and Electronic Health Records (CALIBER) data portal [4], and the Concept Dictionary at the Manitoba Centre for Health Policy [28]. Building web-based concept libraries enables data analysts, researchers, and clinicians to upload and download lists of clinical codes, update previous code lists, and share clinical code data across platforms, which would improve the validation of EHR-based research [22].

Objectives

This study aims to explore the needs of various users, including researchers, clinicians, machine learning experts, and managers, to develop a data portal for phenotypes (a concept library) and to examine why existing concept libraries are not widely used.

Methods

Design

A qualitative study using one-to-one interviews and a focus group discussion was conducted. We recruited a small purposive sample for in-depth one-to-one interviews in the first phase because it allows us to obtain substantial information from a small number of participants while also providing insight into their different viewpoints, needs, and experiences with concept libraries. In the second phase, we recruited a larger sample of participants for the focus group discussion to improve the generalizability of the results. The inclusion criteria were to recruit potential users of concept libraries from various disciplines, including researchers, clinicians, machine learning experts, and managers who conducted studies using routine data generated by data linkage repositories.

For this study, we adopted a semistructured approach. We created semistructured interview questions based on the Krueger and Casey format [29], which included introductory, flow, key, and final questions to be used in one-to-one interviews (Table 1). We also created a list of 10 questions based on the objectives of this study for the focus group session. The purpose of the questions was to generate thoughtful and thorough responses from the participants; therefore, closed-ended questions (eg,

yes or no) were avoided. The interviews and the focus group discussion were audio recorded and transcribed verbatim, and 2 thematic analyses were performed using the 6 steps of Braun and Clarke to identify the themes and subthemes [30].

Table 1. One-to-one interviews' questions guide.

Introductory questions	Follow questions	Key questions	Final questions
To improve repeatable research in Swansea, a team of developers is developing a prototype concept library. This is a portal that allows access to the read codes or International Classification of Diseases–10 codes to identify conditions. Do you think this will be a helpful resource? Is the concept library a good idea that we should continue to develop?	Do you know about other already existing concept libraries? What do you think about them? Something like this exists at UCL[a] called CALIBER[b]. Have you seen CALIBER? Have you used it?	• Do you prefer to use ready-made algorithms or to have access to them to modify them? • In your opinion, how should codes and algorithms be validated, and should they be validated? (Why should or should not?) • There are often different versions of a diagnosis (eg, highly specific and suspected or likely cases). Do you think we need to collect and validate the best two versions of a diagnosis (specific or suspected)? Or do you think we should put all possible methods of identifying a condition, valid or not, and allow the researcher to choose?	• What are your requirements for the concept library for it to be helpful and user-friendly? • What developments would you like to see to improve repeatable research using routine data?

[a]UCL: University College London.

[b]CALIBER: Clinical Disease Research Using Linked Bespoke Studies and Electronic Health Records.

Data Collection

The first author asked 6 participants from a variety of disciplines, including researchers (3/6, 50%), a clinician (1/6, 17%), a machine learning expert (1/6, 17%), and a senior research manager (1/6, 17%), at Swansea University and Cardiff University to participate in one-to-one interviews by email. The invitation email specified the aim and purpose of this study, the duration of each interview (30 minutes), and the location of the interviews, which might be their offices or a convenient and private location on the Swansea University campus.

Semistructured interview questions, which follow the structure proposed by Krueger and Casey [29], were used (Table 1). The structure of the interview questions consisted of introductory, flow, key, and final questions. The purpose of the introductory questions was to help the participants talk freely about their overall experiences. The flow questions were designed to create a smooth transition to the key areas that the authors intended to explore. The final questions were designed to summarize the interview and ensure that the participants did not have further comments [31].

Before conducting the interviews, the first author explained the purpose of the research and what it involved, and at the beginning of each interview, participants received additional verbal and written information about the research project. The interviews were conducted at Swansea University Medical School in a place selected by the participants (eg, their office). After 5 interviews, no new themes were observed and interview 6 confirmed that no new themes emerged. The interviews were audio recorded and transcribed verbatim. Thematic analyses were then performed using the 6 steps of Braun and Clarke to identify the themes and subthemes [30].

All researchers working with the SAIL databank, a national eHealth data linkage infrastructure in Wale (N=34) were invited by email to participate in the focus group discussion, and 14 (14/34, 41%) researchers attended the focus group discussion. In total, 2 focus group discussions, each of which had 7 (7/14, 50%) participants, were held for 2 hours by 2 moderators (ZA and SB), who used the same set of semistructured questions to perform a SWOT (strengths, weaknesses, opportunities, and threats) analysis for the current system for phenotyping and the proposed concept library. We used a SWOT analysis tool in this study because it enabled the participants to discuss what they liked (strengths), what advantages would be gained (opportunities), and what problems (weaknesses) and issues (threats) they felt needed to be tackled. Although the 2 moderators used the same set of questions, the order of the questions was adjusted to the needs of each group.

At the beginning of the focus group discussion, the first author gave a brief presentation about concept libraries, including defining concept libraries, explaining their potential uses, and mentioning examples of some of the existing concept libraries in the United Kingdom. A second presentation about the Swansea University prototype concept library was then given by one of its developers. Feedback from the participants was sought concerning their perceptions of the concept library's needs and their evaluation of the strengths and limitations of the proposed concept library. Participants' perceptions of existing concept libraries, as well as their assessment of the proposed concept library's strengths and limitations, were explored using the following set of semistructured questions:

- What are your thoughts regarding the proposed data portal for phenotypes (a concept library) when it rolls out?
- Do you think this is worth doing? Would you value this?

- Has anybody used existing concept libraries? What have you experienced with them?

Let us talk now about your current system for phenotyping:

- What do you do? What are your methods?
- Are you happy with them? Or what would you like differently?
- What are your thoughts on this plan (building a concept library)?
- Would you use it? Would you share your phenotypes and your phenotyping algorithms?

If you do not want to share your work:

- Can you tell us why? And what motivates you to share it with others?
- Of all the things we have discussed, what is most important to you?
- Is there anything we should have talked about but did not?

The goal of using the SWOT analysis was to identify positive factors that operate together and the potential difficulties that must be identified and solved. During the focus group discussions, participants expressed their own opinions and listened to the opinions of others. As the discussions progressed, participants began to ask questions of one another and share similar experiences. This increased the depth of the conversation. The SWOT analysis gave us a full picture of views and experiences of concept libraries by the participants, making this a holistic evaluation with the ability for participants to hear and comment on each other's responses. Textbox 1 presents a summary of the SWOT analysis in the current system for phenotyping and the proposed concept library. The 2 focus group discussions were audio recorded and transcribed verbatim. Thematic analyses were then conducted using the 6 steps of Braun and Clarke to discover the main themes and subthemes (Table 2).

Textbox 1. A summary of a SWOT (strengths, weaknesses, opportunities, and threats) analysis of the current system for phenotyping and the prototype concept library.

SWOT analysis

Strengths

- Concept libraries provide researchers with a good starting point.
- Publicly available code lists may provide researchers with a history of a particular area of research, such as asthma.
- Referencing previously published lists of codes enables researchers to demonstrate a rationale for using such lists of codes.
- Using research methods developed by others that match the researchers' interests could result in significant time saving.
- Collaboration among researchers is facilitated through sharing and using research methods such as code lists.

Weaknesses

- Searching for and reusing phenotypes and codes is a time-consuming and labor-intensive process.
- There are various lists of codes for each phenotype definition.
- The list of codes chosen by clinicians varies significantly.
- A large number of previously developed code lists could not be repeated.
- Reusing other researchers' data requires programing knowledge such as SQL.
- Some of the ready-made phenotyping algorithms may not be very useful in terms of their general purpose.
- Some existing concept libraries have limited user interfaces.
- Some existing concept libraries are not user-friendly.
- It is unclear who is accountable for the quality of the uploaded codes in concept libraries.
- The validity of the content of concept libraries is unclear.

Opportunities

- Concept libraries must provide user documentation.
- Concept libraries must provide users with training.
- Transparency in sharing the whole approach used to create the code lists is required.
- Establishing a standardized way of defining each specific condition to facilitate comparisons of research outcomes across the United Kingdom.
- Creating a specialized library that stores code lists of a specific condition within a specific set of patients, such as a concept library specializing in chronic conditions in children.
- Creating a concept library that engages a wide variety of users (ie, is easily understandable by clinicians but has some advanced features such as programing skills for more expert users).

Threats

- The inconsistency of data across various databases makes data reuse difficult.
- Lack of confidence in the quality of the list of codes developed by other researchers if they are not cited.
- Access to code lists is limited as some researchers do not publish them alongside their studies.
- Different research outcomes result from a lack of access to a list of codes created by other researchers.
- Data sharing may be inhibited if there are no returns, such as referencing and acknowledgment.
- Concerns about ownership rights discourage data sharing (eg, methods could be used as their own by other researchers before publication).

Table 2. Presentation of the themes and subthemes of the one-to-one interviews.

Themes	Examples of participant narratives
Theme (1): previous opinion of a prototype concept library	
Positive	"If there's a way of doing that already that is set up and is validated and is consistently applied that would be an amazingly useful resource" (researcher 2).
Neutral	"It will be helpful, but it needs to be extended. If they want to build something like this, and it is effectively working as a library, you need two things to be happened: (1) people are happy to feed in their constructs so it builds up, and (2) a useful library, easy to go, to browse, and to borrow phenotypes definitions" (a clinician).
Negative	None
Theme (2): requirements of a prototype concept library	
Usability	
Simplicity	"Simple plain English not in SQL or python" (a clinician).
Searching ability	"What is the type of search engine? Is it a search engine that just does disease phenotypes or also does the health status phenotypes or risk factor phenotypes, symptoms phenotypes?" (a clinician).
Data quality	"It's really just about transparency and documentation. So, anybody can effectively do anything that can be turned into a reproducible research output. The barriers are usually not enough time to comment and document it properly and then not enough quality assurance" (a senior research manager).
Sharing ability	"It would be very useful to share the knowledge about codes such as read codes, ICD 10 codes, or OPCS codes, and share ideas and concepts between other users that will save lots of time" (researcher 3).
Sustainability	
Interoperability	"How interoperable it is with other systems because the major failure of most of these systems is that they're not interoperable, so people don't use them" (a senior research manager).
Accessibility	"So, from a group like myself, or me as a user, we would probably like direct access to the underlying data it stores. So, whether that's through something like SQL directly, or something like that through a statistical package, because where we do lots of bulk type work" (a senior research manager).
Analyzability	"I wanted to look at all health codes of my study population. Then, through machine learning, like feature selection, I tried to identify the most important list of codes, which are associated with the popular health conditions" (researcher 1).
Theme (3): user experience of existing concept libraries	
Aware (used them)	"Yes, so with QOF, we definitely used QOF codes a lot, because obviously going back to the quality assurance question, they'd been assured so that the NHS can use them for remuneration of money and payments. With other systems, we tend to look online to see CALIBER of things with us, then yes we have used outputs from those systems before" (a senior research manager).
Aware (not used them)	"No. I have not used any of these things before so I think there is CALIBER and I think, is that part of what was set up within the previous Farr institute? so I am aware that some of these exist but I haven't looked into them before" (researcher 2).
Not aware	None
Theme (4): user's recommendation to improve repeatable research	"If we want reproducible research, we have to all be using these resources in a similar way or at least we need to be able to understand what previous projects have done. It is about setting things out clearly. Clear definitions, clear sets of codes that people can then either use themselves or build on I think" (researcher 2).

Data Analysis

The interviews and the focus group discussion were analyzed separately following the analysis approach by Braun and Clarke [30]. The transcripts of the interviews and the focus group discussion were read several times, and then the initial codes were grouped into themes and subthemes using a qualitative data analysis software (NVivo, QSR International) [30,32]. ZA had read all the transcripts, and SB read a sample of the transcripts. They independently identified the themes and subthemes, then met regularly to compare them and reach an agreement on what was being done. Themes and subthemes were discussed with respect to their relevance to the research question in the data collected. They critically reviewed the themes again to determine their primary meanings, and similar initial themes were combined into one theme. They discussed the definitions of the relevant themes in the research questions and applied appropriate names to describe each in this study. Textbox 2 provides further description of the thematic analytic steps.

Textbox 2. The 6 thematic analytic steps used for this research.

Thematic analytic steps

Self-familiarizing with the data

- ZA transcribed half of the audio recordings from the interviews (3/6, 50%). The other half of the audio recordings from the interviews (3/6, 50%) and the audio recordings from the focus group discussion were transcribed by professional transcribers. During this phase, ZA read all the interview and focus group discussion transcripts several times, and SB read samples of them. ZA and SB considered all the topics discussed by the participants, recorded notes on these topics in the transcripts, and then organized them in a note book.

Creating initial codes

- After familiarizing themselves with the data, ZA and SB worked independently to identify initial codes from the transcripts that summarized what was said during the interviews and focus group discussion. They organized the identified codes into meaningful groups using qualitative data analysis software (NVivo, QSR International). They used the same coding procedure for all the transcripts.

Searching for themes

- ZA and SB started interpreting the initial codes using their extracted data, and they began grouping the codes with similar meanings together. Using the NVivo software (QSR International), the initial codes were then sorted and labeled into themes and subthemes depending on the meaning or relations shared by the codes.

Revising themes

- ZA and SB critically reviewed and refined themes against the data several times to determine their core meanings, and similar initial themes were combined into one theme. To reach an agreement, themes and subthemes were discussed in terms of their relevance to the research question.

Defining themes

- Each of the themes identified in the previous steps was named and defined by ZA and SB. They used the initial labels created for the themes to provide appropriate names that describe the meaning of the themes in this study. ZA and SB defined each theme based on the content and meaning of their codes, and they examined these definitions in relation to their relevance to the research questions.

Writing up the report

- After defining and naming the themes, ZA and SB began writing the findings for this manuscript. They used quotes from the participants' responses that related to the themes and the research question to illustrate the findings.

Ethics Statement

Ethical approval to conduct the research was approved by the Research Ethics Sub-Committee of Swansea University, project reference number 2019-0007.

Results

Interviews With Users

Overview

In total, 6 one-to-one interviews were conducted, and each interview lasted for approximately half an hour. The analysis of the interviews resulted in 4 main themes, with several subthemes (Table 2). The four main themes are as follows:

1. Previous opinion of a prototype concept library
2. Requirements of a prototype concept library
3. Experience of existing concept libraries
4. Recommendations to improve repeatable research

Previous Opinion of a Prototype Concept Library

The majority of the participants were positive about the prototype concept library and felt that a concept library in principle was a very helpful resource for conducting repeatable research. A machine learning expert mentioned that a concept library will be an extremely useful resource because read codes from general practice and International Classification of Diseases (ICD)–10 codes from hospitals are the most common data items that machine learning experts would like to use most often. They use data linkage repositories to extract the necessary data for machine learning in public health studies, and they use the codes to extract the data from the repositories. Researcher 3 said, "It would be very useful to share the knowledge about codes such as read codes, ICD 10 codes, or OPCS codes, and share ideas and concepts between other users that will save lots of time. It is useful to use verified codes," and researcher 2 stated, "If there's a way of doing that. Already that is set up, and is validated and is consistently applied, that would be an amazingly useful resource."

However, 2 participants (a clinician and a senior research manager in health data science) were not sure about the effectiveness of the prototype concept library because they felt that users had to engage with it for it to be useful and they were not sure how well users would engage: "There is potential that it could be useful as a tool. It will kind of come down to how usable it is, how flexible it is, how well it's maintained, how much of the community uses it" (a senior research manager).

Requirements of the Prototype Concept Library

The participants mentioned several requirements they would like to see in the prototype concept library. For example, they stated that the concept library needed to have high usability. This means that it needs to be simple and easy to use by naïve users: "It should be simple enough, within one or two clicks;

we can find the required data, but also should contain advanced expert features (R, SQL, or Python programing languages) to extract, include, or exclude codes necessary for their studies" (researcher 3) and "Like, in one of my previous projects, I looked at, from a machine learning perspective, I wanted to look at all health codes of my study population. Then, through machine learning, like feature selection, I tried to identify the most important list of codes, which are associated with the popular health conditions" (researcher 1). They also stated that the concept library should have a good search engine so that they can easily find the phenotypes and phenotyping algorithms they want to use. A clinician inquired, "What is the type of search engine you are developing? Is it a search engine that just does disease phenotypes? or also health status phenotypes, risk factor phenotypes, or symptom phenotypes. For example, I am looking for diabetes, but I may also be looking for smoking or alcohol consumption, or symptoms like pain or cough. So, how big is the enterprise and how do you search for what are the appropriate terms? Discussion is needed to know what is it?"

In addition, the participants stated the following requirements:

1. Include the data sources used (eg, codes from general practice, hospital [ICD and Systematized Nomenclature of Medicine], and British National Formulary medication), a general clinical code list for comparison, lists of ontologies along with their variances and versions, and a description of how codes were established: "It is about setting things out clearly. Clear definitions, clear sets of codes that people can then either use themselves or build on, I think" (Researcher 2).
2. Have a clear phenotyping algorithm labeling convention for search engines. A clinician stated,"What do you search on? Thought about what do you call these phenotypes? Is there a consistent in calling them? For example, Type II diabetes, or insulin dependent diabetes" and researcher 1 stated, "So, first of all, for the code reference library, two things are always there in my mind. It's in my opinion again. Number one, they should be validated. Secondly, they should be correctly labelled."
3. Specify why a particular phenotyping algorithm was developed (eg, definite disease or probable/suspected condition definitions): "When I have an algorithm, I want a field that tells me the purpose of the algorithm, a brief description of what the algorithm is intended to do" (a clinician).
4. Illustrate the logic model category used to create phenotyping algorithms (ie, code lists, inclusion or exclusion factors, and clinical or machine learning approach used). "Is this just a code list of inclusion factors? And or exclusion factors? Or is it static? Does it have a tampered relationship? So, some algorithms are present or absence of conditions, some required a tampered dependence. In the logic model categories: Is this a clinically derived algorithm from experts' views or for instance that machine learning derived algorithms" (a clinician).
5. Use ready-made phenotyping algorithms that can be modified to fit the needs of their research. All participants agreed that if they had to create their own phenotyping algorithms because ready-made phenotyping algorithms could not be modified, they needed an easy approach to use a code list in the concept library.

There was an issue regarding how to validate phenotyping algorithms, and most participants expressed their preferences for using all possible methods of identifying a condition, valid or not, to allow the researcher to choose the phenotyping algorithms according to their research requirements: "So, there is no right answer for that because it's going to be very dependent on your research question, your study group, and your study design. So, once again, if the concept tool is going to match multiple different use cases, it's going to need to accommodate for those different types of study design" (a senior research manager). Sharing phenotyping algorithms needed to be easy and not time-consuming, and some felt there needed to be some recognition of their work before they would give their codes. Finally, a concept library must be interoperable with other products or systems: "How interoperable it is with other systems, because the major failure of most of these systems is that they're not interoperable, so people don't use them" (a senior research manager). Most participants wanted the source code (eg, the SQL code for the phenotyping algorithm itself) to be available in a downloadable machine-readable format to be able to access it using specific programing languages such as R, SQL, or Python.

Experience of Existing Concept Libraries

All participants stated that they were aware of some existing concept libraries, such as CALIBER and ClinicalCodes.Org (both in the United Kingdom), but most of them did not use them. The reasons given for not using them were that they already had their own self-made concept libraries (eg, concepts they have used before) or the available concept libraries did not provide phenotyping algorithms that fit their studies. For example, a machine learning expert mentioned the reasons for not using two of the existing concept libraries, namely the Concept Dictionary at the Manitoba Centre for Health Policy in Canada and CALIBER in the United Kingdom were, "Canadian systems provide Canadian data for their studies, CALIBER is specific for cardiovascular disease and does not have many concepts in it." Conversely, 2 of the participants mentioned that they used some existing concept libraries to extract and develop phenotyping algorithms for their studies: "We definitely used QOF codes a lot, with other systems, we tend to look online to see CALIBER, we have used outputs from those systems before" (a senior research manager).

Recommendations to Improve Repeatable Research

The participants suggested the following recommendations to improve repeatable research output using routine data:

1. There should be a drive for more transparency in research methods documentation, such as publishing complete phenotype definitions and clear code lists. A senior research manager stated, "It's really just about transparency and documentation. So, anybody can effectively do anything that can be turned into a reproducible research output," and researcher 2 said, "If we want reproducible research, we have to all be using these resources in a similar way or at least we need to be able to understand what previous

projects have done. It is about setting things out clearly. Clear definitions, clear sets of codes that people can then either use themselves or build on, I think."

2. Providing opportunities for researchers to collaborate rather than working in isolation, "The barriers are usually not enough time to comment and document it properly and then not enough quality assurance. So, if there was more time and or more availability of those kinds of opportunities for people to collaborate rather than doing things in isolation, there's almost all the research we do here could be turned into a reproducible type of output" (a senior research manager).
3. Develop a concept library that enables researchers to begin classifying population outcomes using uniform codes: "I think that a resource like this is a very good step in the right direction because I think what people need to start doing is using consistent codes in order to identify conditions or outcomes within populations" (researcher 2).
4. Provide validated phenotyping algorithms that researchers can use directly to avoid duplication, with the ability to modify them to meet their own research needs: "For each project, it always has some specific requirement which is unique, which is not common. There are some things which are common, and there are a few things which are very unique. So, we need to have some algorithms which we can just use to, you know, just to avoid the duplication, but also, we need to have control of the algorithms, so that we know only that these bits are going to be different for this project, so I'm going to replace, change, modify this bit, and we'll run it" (researcher 1).

Focus Group Discussions

Overview

Of the 34 invited researchers, 14 (41%) attended the focus group discussion. These participants were researchers (14/34, 41%) from Swansea University who were working with the SAIL data in the Data Science Building. Of the 14 participants, 5 (36%) were female participants and 9 (64%) were male participants. Furthermore, 6 (43%) participants were PhD holders, 6 (43%) were Master's degree holders, and 2 (14%) were Bachelor's degree holders (Table 3).

Table 3. A summary of general information on the participants in the focus group discussions (N=14).

Parameters	Information
Current job position, n (%)	• Data scientist, 13 (93) • Financial planner, 1 (7)
Sex, n (%)	• Female, 5 (36) • Male, 9 (64)
Education, n (%)	• PhD degree, 6 (43) • Master's degree, 6 (43) • Bachelor's degree, 2 (14)

Research interests	• Data scientists
	  • Concept libraries
	  • Repeatable research with large health data
	  • Phenotyping and code lists of cancer disease
	  • Respiratory disease
	  • Algorithm or reusable codes development
	  • Asthma
	  • Collaboration in research methods
	  • Data analysis
	  • Machine learning
	  • Arthritis
	  • Health informatics
	  • Musculoskeletal disorders
	  • Healthy aging
	  • Gut—brain axis
	  • Neurodegenerative conditions
	  • Statistical methods
	  • Epidemiology
	  • Cancer
	• Financial planners
	  • Intervention between primary care and secondary care and how they interact

The focus group discussion was held for 2 hours to perform a SWOT analysis of the current system for phenotyping and the proposed concept library, which was recorded and transcribed, and thematic analysis was conducted on the transcripts, which resulted in the identification of the following seven main themes:

1. Facilitators for and barriers to participants' contributing their research methods
2. Facilitators for and barriers to participants' use of other researchers' methods
3. Participants' concerns about the prototype concept library

4. The requirements of the participants for the prototype concept library
5. Participants' recommendations to improve repeatable research
6. Participants' perceptions of their current phenotyping system
7. Participants' use and perceptions of existing concept librariess

Facilitators and Barriers to Participants' Contributing Their Research Methods

Facilitators

Several facilitators were identified by participants as motivators for them to share their work (eg, phenotyping algorithms and code lists). Many participants stated that being credited appropriately (eg, receiving citations from other researchers) would motivate them to share their work: "If whoever's using it acknowledges it's use in whatever they publish, at least you're getting some recognition" (data scientist 8) and "If there were DOIs attached to the code list of algorithms, when people are publishing, there's an incentive for putting it on there, because they're able to demonstrate the impact their work has had" (data scientist 4).

Some participants stated that communicating with their research team would encourage them to organize team resources and discuss research findings from other researchers who used their code lists. However, improving research opportunities, increasing academic achievement, and sharing knowledge through collaboration with other researchers working in the same organization would motivate some of the participants to share their work: "I think there's benefit to the organization, and there has to be benefit to the people contributing to it" (data scientist 4). In general, researchers work in an organization (eg, a university or a research institute), and they work hard to improve the research outcomes of their organization. Some participants stated that advancing the research base and saving other researchers' time and effort would stimulate them to share their work: "Surely if you've done something you think really worthwhile, you want other people to use it, as well, because then that furthers the research" (data scientist 6).

Barriers

On the other hand, the participants pointed out several barriers that could inhibit them from sharing their work (eg, phenotyping algorithms and code lists) with other researchers. Some participants argued that it is easy to build a phenotyping algorithm that fits exactly their needs, but it is more challenging to develop a general one, so it can be used by others (eg, many clinical researchers have created phenotyping algorithms for particular research, and these algorithms are difficult to generalize).

Several participants mentioned that a lack of return for their hard work (eg, not receiving any credit from others, such as referencing when they reuse their data) would prevent them from sharing their work: "How do you enforce that people are going to give you credit? It doesn't happen sometimes, when referencing, saying where they got it from. You've just got to hope they do" (data scientist 11). Some participants were worried about their intellectual rights (eg, if they shared their methods such as phenotyping algorithms before publication, other researchers would use them as their own).

Facilitators and Barriers to Participants' Use of Other Researchers' Research Methods

Facilitators

The participants mentioned several facilitators that would encourage them to reuse research methods developed by others, such as the following:

1. Using existing code lists can save them a lot of time and effort, which they frequently spend creating new code lists from scratch: "It's the first stage of every single process, and we tend to get two or three months of work, until we get to that final code list, and we can now start looking at the cases" (data scientist 10).
2. Reusing available data, such as code lists, is a good place to start for researchers (for example, they can use them to examine new ideas and gain new insights):"Having code lists would be such a help, to get you started. They always want things like BMI and weight and height. There are hundreds of codes for those. The smoking codes, having a list, even if you don't use the algorithm that they've developed, is a huge bonus" (data scientist 12).
3. Using the work of others as a reference to compare research outcomes, and researchers want to prove that there is a basis for the use of such codes.

Barriers

Conversely, the participants pointed out several barriers that could inhibit them from reusing methods developed by other researchers such as the following:

1. Poor data quality discourages researchers from reusing it: "You could upload complete garbage" (data scientist 1).
2. Some phenotyping algorithms will not work outside the population in which they were developed. For example, code developed in Canada may not be relevant to finding conditions in general practitioner data in the United Kingdom: "Yes, it works in their population, because where they've trained it." (data scientist 5). .
3. Whether the data are useful to researchers plays an important role in the decision to reuse them (eg, researchers would not use a phenotyping algorithm if its general purpose did not match their interests): "Yes, a general-purpose algorithm may or may not be very useful to have it to see what they've done, but you may not use it" (data scientist 12).

Participants' Concerns About the Prototype Concept Library

When researchers decide where to deposit, share, and reuse data, they prefer to use approved concept libraries: "Is it going to be approved?" (a financial planner). Moreover, some participants stated that it is not clear who is responsible for the quality of the phenotyping algorithm, if this is the responsibility of the developers running the concept library or the responsibility of the researchers uploading the phenotyping algorithms: "If people send the codes, the onus of the quality

of that code list you would still want to be the responsibility of the researcher to be submitting worthwhile codes. You don't want to then be the guardian of the quality of the code list. You still need to know where the responsibilities lie" (data scientist 4). Researchers do not want to upload phenotyping algorithms if they could be *blamed* for flaws, and health informatic developers do not want to take responsibility for the phenotyping algorithms that were uploaded.

The participants expressed their concerns about the completeness rate of the phenotyping algorithms. They would like to know the percentage of the gap to be considered when using a phenotyping algorithm from the prototype concept library: "What is the completeness rate? For certain things, we know there are gaps. If the gap is 20%, is that something I should be including in any algorithm I'm considering?" (data scientist 8). In addition, there has been a question as to whether codes need to be peer reviewed so that quality is evaluated.

Requirements of the Participants for the Prototype Concept Library

Usability

1. Learnability: Some participants said they would like the concept library to be easily understandable by clinicians, who acknowledge the clinical definition of the code lists with little technical skills to simply point and click the selected code lists, whereas other participants requested the availability of advanced functions to be used by expert users: "The concept library should be easy. Someone needs to train us" (data scientist 9).
2. User documentation: A collection of well-defined task-oriented documentation for users was required by some participants. They want a user documentation that consists of clear, step-by-step instructions on how to use the concept library and gives examples of what the user can see at each step (eg, screenshots would be useful): "Concept library should have some documentation" (data scientist 9).
3. Data quality: Some participants required the availability of a consistent method for identifying each specific condition to ensure that what researchers are doing is compatible within their immediate team but also within the broader research community in the United Kingdom to facilitate a comparison of research outcomes. Other participants stated that they needed a predefined list and a uniform approach describing how to use existing codes of additional diagnoses, such as smoking: "Additional things like smoking and alcohol status are used a lot, but they're usually very different for every project. We should have a more uniform way of doing it, like, we'll take that bit off the shelf and use it, and do the bespoke bit for things that need to be bespoke" (data scientist 5). If there are multiple code lists for the same condition, some participants proposed that versions be generated to describe each particular condition: "So, it would be relevant that there were multiple lists for the same condition, if you've got a version and way of defining a certain condition" (data scientist 4).
4. Transparency: Several participants required transparency in sharing the entire approach used in developing the code lists, including phenotyping algorithms and the methods used. They stated that if they use a code list for each comorbidity of a condition, they will build an entirely different score over the years. Therefore, transparency in the documentation of research methods would help them to know which score is the best.

Sustainability

1. Accessibility: Several participants needed the availability of an access control that allows access to the codes only after publication, while at the first stage of the study, researchers spent a lot of time and effort developing them, and they feared someone else could publish work faster than them using the algorithms: "There should be an option in the concept library for lists that have been published. People can develop them, but if they're not published, you don't have to use them" (data scientist 3).
2. Licensing: Some participants needed to know which type of license was adopted by the developers of the concept library (eg, researchers can have one that means any researcher can take it and use it, or they can have one that means researchers can use it but not for commercial purposes).
3. User community: Several participants required users to quote a reference if publishing papers based on the results (partially or completely) derived from the concept library: "If I want to use someone else's work, I think that's the norm, and should be in this economy. Anything, not just code. To use this, I should reference that it's based on this or other thing completely, or a part of it" (data scientist 2). Referencing helps to determine whether there is or will be an active user community for the concept library and the codes used: "It potentially would make your publication more discoverable. If there's a whole community of users using this" (data scientist 1).

Participants' Recommendations to Improve Repeatable Research

Of the 14 participants, 9 (64%) suggested that the prototype concept library should be accessible both in the United Kingdom and globally and practically available to enable researchers around the world to use an web-based secure platform, which stores codes and other logic, and to encourage researchers to contribute their codes to promote research: "Should be open for the United Kingdom" (data scientist 9). However, a participant recommended that the prototype concept library should be closed at the beginning to ensure it is working and then to become opened as researchers build trust: "You might need to restrict it, to start with, to make sure it works. Otherwise, everyone will see the problems you might have" (data scientist 12). In addition to know who is using the concept library, data scientist 8 suggested that it should have request sharing followed by open sharing.

Accessibility to research data has significant potential for scientific advancement as it promotes the replication of research results and enables the use of old data in new contexts. With respect to this, some participants suggested that funders and publishers should obligate researchers to share their research data such as code lists: "Some sort of obligation by funders to

share this" (data scientist 2) and "Publishers, as well" (data scientist 8).

A participant suggested the use of preauthorization of publication by journals based on the research protocol because researchers can put their protocol first, and all the limitations are actually corrected before they run the research. This approach has many advantages for both the researcher and the publisher, as it improves the quality of the output. Another participant recommended the creation of a discussion forum in the concept library to facilitate collaboration among researchers on just about any topic (eg, they can share their ideas, submit their comments, and discover new ideas): "Make it almost a forum" (data scientist 8).

Participants' Perceptions of Their Current Phenotyping System

The participants mentioned several problems associated with the current phenotyping system. For example, they have to search for codes from different databases, which use different coding systems such as read codes and ICD-10 codes, and then they have to validate the selected code lists with experts in the field such as clinicians: "I have to google all of this and search what was there within the community. I have to go to CALIBER, I have to go to Manchester, or there is a work in Edinburgh University, do some work there. Do the search. I have to go there, see the ability to work, and start. It does take a lot of time. Based on my study of Google, I have to start a record, and I have to validate it, verify with other people, clinicians or researchers. It's a long process" (data scientist 9).

Although they could find some codes on the web, they still had to locate the list manually, copy it, and enter the codes into their scripts. Often, they might spend a few days on it, and they might miss obscure codes or even use irrelevant codes: "Starting from scratch, I would go online to see what's available. Go into other people's and see their code lists" (data scientist 11). With respect to this, some participants said that they preferred to use code lists that were referenced or used by other researchers.

Some participants reported that the read code lists chosen by the researchers were different from the read code lists chosen by general practitioners. For example, they found that there were some very clear codes, but they were rarely used by general practitioners: "What we get in the read code list isn't necessarily what the GPs are recording it under" (data scientist 12). They also stated that there is a significant difference between what one general practitioner may say in a list of codes versus another: "For example, there is no single entity code for asthma. There are different entities. If you want to find specific things within asthma, there's a list of codes for them" (data scientist 2).

Participants' Use and Perceptions of Existing Concept Libraries

Not all participants had previously used some of the existing concept libraries. However, most of those who used some of them expressed negative perceptions. For example, several participants stated that the concept libraries they used were not user-friendly (ie, they were difficult to use by new users): "For CALIBER, it seems not so user friendly. It's not easy. You have to know first. Someone needs to train you up. For new users, it's difficult to get inside CALIBER. The concept library should be easy. Someone needs to train us. Concept library should have some documentation" (data scientist 9). Therefore, training and good user documentation are required. A further problem for some participants was the inconsistency of data among various databases, which makes reuse of data quite challenging. "But if there is something that gets secondary and primary care involved, and there's a registry, if the definitions that are created in Manchester, how easy will it be to apply it to, for example, in Wales or Scotland, where registry is a bit different?" (data scientist 8).

Participants who did not use any of the existing concept libraries expressed different perceptions about them. For example, some participants reported that they wanted to explore available concept libraries. Others, however, expressed doubts about the quality and validity of the data stored in these concept libraries, which could prevent them from using them: "I haven't looked at them myself, but if you go on this clinical code site and you type in diabetes, there are 50 different code lists people have put together for diabetes" (data scientist 6). Some participants stated that the main reason for not using any of the existing concept libraries is not finding a concept library that matches their studies. The developers of concept libraries may consider building a specialized library that stores code lists of a particular condition within a specific group of patients according to researchers' needs, such as developing a concept library that specializes in chronic conditions in children.

Discussion

Principal Findings

Development of a concept library that meets users' expectations is extremely useful for repeatable research (eg, researchers would be able to use archived code lists to compare studies). This study found that, although in principle, everyone felt that a digital portal containing a concept library would be very helpful, there were many requirements needed before its development. It needs to engage a wide variety of users if it is to be used (and current concept libraries are not widely used), which means that it has to be very simple (point and click) for some, but it should have the software and usability to manipulate and design phenotyping algorithms for more advanced users. In addition, it needs to have a very high-quality search engine so that it is very easy to find information, and for it to expand, there needs to be a reason for users to upload their phenotyping algorithms, which need to be very easy and quick.

This study indicated that although most of the interviewees expressed positive impressions about the idea of building a prototype concept library, approximately half of the participants expressed an interest in contributing to it. For the prototype concept library to work, researchers must engage with it and upload their codes there so that other people can use them. If researchers did not share their codes in the prototype concept library, this would usually mean an empty library. For better adoption of the prototype concept library, it is recommended that the developers consider the various facilitators for and

barriers to participants sharing their work and reusing the work of others.

The findings of the focus group discussion demonstrate that facilitators for the participants' sharing of their research methods vary across four categories: (1) personal drivers (eg, obtaining appropriate credit, such as citations)—this confirms the results of earlier studies that suggest that researchers may be motivated to share their work if sharing leads to an increase in their citations [33-35], (2) benefits for their research team (eg, sharing information to promote research within their team) [36,37], (3) benefits for their organization (eg, collaboration among researchers working within the same organization would advance their organization's research outcomes), and (4) benefits for the research community (eg, expanding research base) [38]. With respect to this, Cragin et al [39] have stated, "As a research group gets larger and more formally connected to other research groups, it begins to function more like big science."

There were several barriers that could inhibit the participants from sharing their research methods, such as the expected performance of the shared methods (eg, they felt that building a general phenotyping algorithm to be used by others is very difficult) [40] and lack of personal benefits such as recognition (eg, they were worried about not being referenced by researchers who used their methods). In relation to this, Molloy et al [41]. reported that researchers can be discouraged from sharing their work by fear of not obtaining sufficient credit. Therefore, a safeguard against uncredited use is necessary [42]. In addition, participants mentioned that they were afraid that their methods would be used by other researchers as their own before publication. The results of the study conducted by Huang et al [43] indicated that although most participants are interested in sharing papers related to biodiversity data, >60% of the participants were reluctant to share primary data before publication. Moreover, findings from this study correspond with other studies regarding the need to adapt impact metrics to promote data sharing [44,45] because researchers would not be able to measure the success of their methods if metrics are not available. Unless these obstacles are resolved, the sharing of data in concept libraries is unlikely to increase significantly.

Several facilitators encouraged participants to reuse research methods developed by others. They reported that reusing code lists created by other researchers would make their task much easier, save them a lot of time, and help to demonstrate that there is a justification for using such codes. These findings are consistent with those of the previous studies. For example, Anneke and Helen reported that researchers are using open research data to "be aware of the state of the art and not recreate the wheel, as well as access to more data and generating fresh insights" [46].

The results of this study indicate that more than half of the participants were not satisfied with their current system for phenotyping for several reasons, including the lack of accessibility of other researchers' work, such as code lists, which could affect research outcomes and the fact that reusing publicly available code lists consumes a lot of time and requires lots of work [38]; lack of confidence in web-based code lists if they are not cited by other researchers; lack of availability of a consistent approach for defining covariates such as smoking; and the selected read code lists by the researchers are different from the selected read code lists by the general practitioners. It seems that their current approach lacks confidence and is time-consuming and effort-intensive.

This study demonstrates that existing concept libraries are not widely used, and most participants who used some of the existing concept libraries expressed negative impressions about them (eg, they do not provide training or user documentation, and they are difficult to use) [36-38]. Lack of knowledge of the existence of concept libraries and how to use them is generally described as an obstacle to data sharing [47]. As existing concept libraries are not used by all researchers, obstacles that inhibit researchers from using them need to be addressed when building new concept libraries.

Strengths and Limitations

To our knowledge, this is the first study aimed at identifying the needs of various users of a concept library. The findings of this study would have a significant impact on improving the efficiency of existing concept libraries by informing their developers about the different requirements, facilitators, barriers, and recommendations of the various users. In addition, this work will greatly inform the developers of new concept libraries to improve access to and collaboration with EHRs' routine data, which is part of an all-UK agenda, and the findings of this study will have implications for other countries working to access and share EHRs' routine data.

This study has some limitations that should be addressed in future studies. The first limitation is that we had a time limit on how long we could talk to the participants because each one-to-one interview was given 30 minutes. As a result, the number of questions we could ask and the amount of time we could spend on each question were limited. The second limitation is that all the participants of the interviews and focus group discussion were recruited because they used the SAIL databank, a national eHealth data linkage infrastructure in Wale, so they mostly talked about the Swansea concept library in the SAIL databank. As the discussion focused on the SAIL databank, its generalization to other concept libraries was limited.

Conclusions

In conclusion, although it may seem beneficial for researchers to reuse methods developed by others, such as code lists, some researchers who created them prefer not to share them because they worked hard to create them and would rather publish them first to ensure their academic rights, such as being referenced [48]. The major challenge is that some researchers would like to use the work of other researchers, but they do not want to contribute their work to concept libraries. Open sharing can be more difficult in the research community as researchers compete for grants, work promotions, and publication quotations [48]. They think carefully about how, when, and where to share their work as they have spent a vast amount of time and effort to develop it [47]. A solution to these issues would be to encourage researchers to contribute data to the prototype concept library in such a way that the shared data is understandable and reusable

(eg, ensuring uploading of adequate documentation) for the public good rather than for personal gains.

Acknowledgments

This research was supported by the Kuwait Cultural Office in London, HDRUK (Health Data Research UK), and the National Centre for Population Health and Wellbeing.

Conflicts of Interest

None declared.

References

1. Badawi O, Brennan T, Celi LA, Feng M, Ghassemi M, Ippolito A, MIT Critical Data Conference 2014 Organizing Committee. Making big data useful for health care: a summary of the inaugural mit critical data conference. JMIR Med Inform 2014 Aug 22;2(2):e22 [FREE Full text] [doi: 10.2196/medinform.3447] [Medline: 25600172]
2. Wei W, Teixeira P, Mo H, Cronin R, Warner J, Denny J. Combining billing codes, clinical notes, and medications from electronic health records provides superior phenotyping performance. J Am Med Inform Assoc 2016 Apr;23(e1):20-27 [FREE Full text] [doi: 10.1093/jamia/ocv130] [Medline: 26338219]
3. Hripcsak G, Albers D. Next-generation phenotyping of electronic health records. J Am Med Inform Assoc 2013 Jan 01;20(1):117-121 [FREE Full text] [doi: 10.1136/amiajnl-2012-001145] [Medline: 22955496]
4. Morley K, Wallace J, Denaxas S, Hunter R, Patel R, Perel P, et al. Defining disease phenotypes using national linked electronic health records: a case study of atrial fibrillation. PLoS One 2014;9(11):e110900 [FREE Full text] [doi: 10.1371/journal.pone.0110900] [Medline: 25369203]
5. 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. JMIR Med Inform 2021 Aug 03;9(8):e24405 [FREE Full text] [doi: 10.2196/24405] [Medline: 34342589]
6. Schleyer T, Song M, Gilbert G, Rindal B, Fellows J, Gordan VV, et al. Electronic dental record use and clinical information management patterns among practitioner-investigators in The Dental Practice-Based Research Network. J Am Dent Assoc 2013 Jan;144(1):49-58 [FREE Full text] [doi: 10.14219/jada.archive.2013.0013] [Medline: 23283926]
7. Wang S. Opportunities and challenges of clinical research in the big-data era: from RCT to BCT. J Thorac Dis 2013 Dec;5(6):721-723 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2013.06.24] [Medline: 24409345]
8. Pendergrass S, Crawford D. Using electronic health records to generate phenotypes for research. Curr Protoc Hum Genet 2019 Jan;100(1):e80 [FREE Full text] [doi: 10.1002/cphg.80] [Medline: 30516347]
9. Kim HH, Kim B, Joo S, Shin S, Cha HS, Park YR. Why do data users say health care data are difficult to use? A cross-sectional survey study. J Med Internet Res 2019 Aug 06;21(8):e14126 [FREE Full text] [doi: 10.2196/14126] [Medline: 31389335]
10. Banda JM, Seneviratne M, Hernandez-Boussard T, Shah NH. Advances in electronic phenotyping: from rule-based definitions to machine learning models. Annu Rev Biomed Data Sci 2018 Jul 20;1(1):53-68 [FREE Full text] [doi: 10.1146/annurev-biodatasci-080917-013315] [Medline: 31218278]
11. Pathak J, Kho AN, Denny JC. Electronic health records-driven phenotyping: challenges, recent advances, and perspectives. J Am Med Inform Assoc 2013 Dec 01;20(e2):206-211 [FREE Full text] [doi: 10.1136/amiajnl-2013-002428] [Medline: 24302669]
12. Richesson RL, Sun J, Pathak J, Kho AN, Denny JC. Clinical phenotyping in selected national networks: demonstrating the need for high-throughput, portable, and computational methods. Artif Intell Med 2016 Jul;71:57-61 [FREE Full text] [doi: 10.1016/j.artmed.2016.05.005] [Medline: 27506131]
13. Vezyridis P, Timmons S. Evolution of primary care databases in UK: a scientometric analysis of research output. BMJ Open 2016 Oct 11;6(10):e012785 [FREE Full text] [doi: 10.1136/bmjopen-2016-012785] [Medline: 27729352]
14. Herrett E, Gallagher A, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol 2015 Jun;44(3):827-836 [FREE Full text] [doi: 10.1093/ije/dyv098] [Medline: 26050254]
15. Blak B, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Inform Prim Care 2011 Jul 01;19(4):251-255 [FREE Full text] [doi: 10.14236/jhi.v19i4.820] [Medline: 22828580]
16. Hippisley-Cox J, Stables D, Pringle M. QRESEARCH: a new general practice database for research. Inform Prim Care 2004 Feb 01;12(1):49-50. [doi: 10.14236/jhi.v12i1.108] [Medline: 15140353]
17. Ford DV, Jones KH, Verplancke JP, Lyons RA, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. BMC Health Serv Res 2009 Sep 04;9:157 [FREE Full text] [doi: 10.1186/1472-6963-9-157] [Medline: 19732426]

18. Al Sallakh MA, Vasileiou E, Rodgers S, Lyons R, Sheikh A, Davies G. Defining asthma and assessing asthma outcomes using electronic health record data: a systematic scoping review. Eur Respir J 2017 Jun;49(6) [FREE Full text] [doi: 10.1183/13993003.00204-2017] [Medline: 28619959]
19. Lu M, Chacra W, Rabin D, Rupp LB, Trudeau S, Li J, et al. Validity of an automated algorithm using diagnosis and procedure codes to identify decompensated cirrhosis using electronic health records. CLEP 2017 Jul;9:369-376 [FREE Full text] [doi: 10.2147/clep.s136134]
20. Manuel D, Rosella LS, Stukel TA. Importance of accurately identifying disease in studies using electronic health records. Br Med J 2010 Aug 19;341:c4226 [FREE Full text] [doi: 10.1136/bmj.c4226] [Medline: 20724404]
21. Nicholson A, Tate A, Koeling RC, Cassell JA. What does validation of cases in electronic record databases mean? The potential contribution of free text. Pharmacoepidemiol Drug Saf 2011 Mar;20(3):321-324 [FREE Full text] [doi: 10.1002/pds.2086] [Medline: 21351316]
22. Springate D, Kontopantelis E, Ashcroft D, Olier I, Parisi R, Chamapiwa E, et al. ClinicalCodes: an online clinical codes repository to improve the validity and reproducibility of research using electronic medical records. PLoS One 2014;9(6):e99825 [FREE Full text] [doi: 10.1371/journal.pone.0099825] [Medline: 24941260]
23. Bhattarai N, Charlton J, Rudisill C, Gulliford M. Coding, recording and incidence of different forms of coronary heart disease in primary care. PLoS One 2012;7(1):e29776 [FREE Full text] [doi: 10.1371/journal.pone.0029776] [Medline: 22276128]
24. Gulliford M, Charlton J, Ashworth M, Rudd A, Toschke A, eCRT Research Team. Selection of medical diagnostic codes for analysis of electronic patient records. Application to stroke in a primary care database. PLoS One 2009 Sep 24;4(9):e7168 [FREE Full text] [doi: 10.1371/journal.pone.0007168] [Medline: 19777060]
25. Sargeant J, O'Connor AM, Dohoo I, Erb H, Cevallos M, Egger M, et al. Methods and processes of developing the Strengthening the Reporting of Observational Studies in Epidemiology - Veterinary (STROBE-Vet) statement. J Vet Intern Med 2016 Nov;30(6):1887-1895 [FREE Full text] [doi: 10.1111/jvim.14574] [Medline: 27859753]
26. Harron K, Benchimol E, Langan S. Using the RECORD guidelines to improve transparent reporting of studies based on routinely collected data. Int J Popul Data Sci 2018 Jan 10;3(1):2 [FREE Full text] [doi: 10.23889/ijpds.v3i1.419] [Medline: 30542668]
27. Benchimol E, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 2015 Oct;12(10):e1001885 [FREE Full text] [doi: 10.1371/journal.pmed.1001885] [Medline: 26440803]
28. Smith M, Turner K, Bond R, Kawakami T, Roos LL. The concept dictionary and glossary at MCHP: tools and techniques to support a population research data repository. Int J Popul Data Sci 2019 Dec 05;4(1):1124 [FREE Full text] [doi: 10.23889/ijpds.v4i1.1124] [Medline: 32935033]
29. McQuarrie E, Krueger R. Focus groups: a practical guide for applied research. J Mark Res 1989 Aug;26(3):371 [FREE Full text] [doi: 10.2307/3172912]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
31. Onwuegbuzie A, Dickinson W, Leech N, Zoran A. A qualitative framework for collecting and analyzing data in focus group research. Int J Qual Methods 2009 Sep 01;8(3):1-21 [FREE Full text] [doi: 10.1177/160940690900800301]
32. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners. Thousand Oaks, CA: Sage Publications; 2013.
33. Patel D. Research data management: a conceptual framework. Libr Rev 2016 Jul 04;65(4/5):226-241 [FREE Full text] [doi: 10.1108/lr-01-2016-0001]
34. Piwowar H, Vision T. Data reuse and the open data citation advantage. PeerJ PrePrints 2013;1:1-25 [FREE Full text] [doi: 10.7287/peerj.preprints.1v1]
35. Viseur R. Open science: practical issues in open research data. In: Proceedings of 4th International Conference on Data Management Technologies and Applications - DATA. 2015 Presented at: 4th International Conference on Data Management Technologies and Applications - DATA; 2015; Colmar, Alsace, France p. 201-206. [doi: 10.5220/0005558802010206]
36. Childs S, McLeod J, Lomas E, Cook G. Opening research data: issues and opportunities. Rec Manag J Internet 2014;24(2):142-162 [FREE Full text] [doi: 10.1108/rmj-01-2014-0005]
37. Dai S, Li H, Xiong J, Ma J, Guo H, Xiao X, et al. Assessing the extent and impact of online data sharing in eddy covariance flux research. J Geophys Res Biogeosci 2018 Jan 15;123(1):129-137 [FREE Full text] [doi: 10.1002/2017jg004277]
38. de Almeida UB, Fraga B, Giommi P, Sahakyan N, Gasparyan S, Brandt C. Long-term multi-band and polarimetric view of Mkn 421: motivations for an integrated open-data platform for blazar optical polarimetry. Galaxies 2017 Nov 30;5(4):90 [FREE Full text] [doi: 10.3390/galaxies5040090]
39. Cragin M, Palmer C, Carlson J, Witt M. Data sharing, small science and institutional repositories. Philos Trans A Math Phys Eng Sci 2010 Sep 13;368(1926):4023-4038 [FREE Full text] [doi: 10.1098/rsta.2010.0165] [Medline: 20679120]
40. Ceci SJ. Scientists' attitudes toward data sharing. Sci Technol Hum Val 2018 Mar 01;13(1-2):45-52 [FREE Full text] [doi: 10.1177/016224398801300106]

41. Molloy J. The open knowledge foundation: open data means better science. PLoS Biol 2011 Dec 6;9(12):e1001195 [FREE Full text] [doi: 10.1371/journal.pbio.1001195] [Medline: 22162946]
42. Ostell J. Data sharing: standards for bioinformatic cross-talk. Hum Mutat 2009 Apr;30(4):vii [FREE Full text] [doi: 10.1002/humu.21013]
43. Huang X, Hawkins B, Lei F, Miller G, Favret C, Zhang R. Willing or unwilling to share primary biodiversity data: results and implications of an international survey. Conserv Lett 2012;5(5):399-406 [FREE Full text] [doi: 10.1111/j.1755-263x.2012.00259.x]
44. Costello MJ. Motivating online publication of data. BioScience 2009 May;59(5):418-427. [doi: 10.1525/bio.2009.59.5.9]
45. Parr C. Open sourcing ecological data. Bioscience 2007;57(4):309-310 [FREE Full text] [doi: 10.1641/b570402]
46. Zuiderwijk A, Spiers H. Sharing and re-using open data: a case study of motivations in astrophysics. Int J Inf Manag 2019 Dec;49:228-241 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2019.05.024]
47. Fecher B, Friesike S, Hebing M. What drives academic data sharing? PLoS One 2015;10(2):e0118053 [FREE Full text] [doi: 10.1371/journal.pone.0118053] [Medline: 25714752]
48. Breeze J, Poline J, Kennedy D. Data sharing and publishing in the field of neuroimaging. GigaSci 2012 Jul 12;1(1):2 [FREE Full text] [doi: 10.1186/2047-217x-1-9]

Abbreviations

EHR: electronic health record
CALIBER: Clinical Disease Research Using Linked Bespoke Studies and Electronic Health Records
ICD: International Classification of Diseases
SAIL: Secured Anonymized Information Linkage
SWOT: strengths, weaknesses, opportunities, and threats

Edited by A Kushniruk; submitted 07.06.21; peer-reviewed by L Lix, Y Tani, K Ludlow; comments to author 26.08.21; revised version received 17.11.21; accepted 05.12.21; published 15.03.22.

Please cite as:
Almowil Z, Zhou SM, Brophy S, Croxall J
Concept Libraries for Repeatable and Reusable Research: Qualitative Study Exploring the Needs of Users
JMIR Hum Factors 2022;9(1):e31021
URL: https://humanfactors.jmir.org/2022/1/e31021
doi:10.2196/31021
PMID:35289755

Original Paper

Acceptance of the Use of Artificial Intelligence in Medicine Among Japan's Doctors and the Public: A Questionnaire Survey

Honoka Tamori[1*], MHS; Hiroko Yamashina[2,3*], PhD; Masami Mukai[1,4]; Yasuhiro Morii[5], PhD, MBA; Teppei Suzuki[2,6], PhD, MBA; Katsuhiko Ogasawara[2], PhD, MBA

[1]Graduate School of Health Sciences, Hokkaido University, Sapporo, Japan

[2]Faculty of Health Sciences, Hokkaido University, Sapporo, Japan

[3]Fukushima Medical University, Fukushima, Japan

[4]Division of Medical Informatics, National Cancer Center Hospital, Chuo, Japan

[5]Center for Outcomes Research and Economic Evaluation for Health, National Institute of Public Health, Wako, Japan

[6]Hokkaido University of Education, Iwamizawa Campus, Iwamizawa, Japan

[*]these authors contributed equally

Corresponding Author:
Katsuhiko Ogasawara, PhD, MBA
Faculty of Health Sciences
Hokkaido University
N12-W5, Kita-ku
Sapporo, 0600812
Japan
Phone: 81 11 706 3409
Email: oga@hs.hokudai.ac.jp

Abstract

Background: The use of artificial intelligence (AI) in the medical industry promises many benefits, so AI has been introduced to medical practice primarily in developed countries. In Japan, the government is preparing for the rollout of AI in the medical industry. This rollout depends on doctors and the public accepting the technology. Therefore it is necessary to consider acceptance among doctors and among the public. However, little is known about the acceptance of AI in medicine in Japan.

Objective: This study aimed to obtain detailed data on the acceptance of AI in medicine by comparing the acceptance among Japanese doctors with that among the Japanese public.

Methods: We conducted an online survey, and the responses of doctors and members of the public were compared. AI in medicine was defined as the use of AI to determine diagnosis and treatment without requiring a doctor. A questionnaire was prepared referred to as the unified theory of acceptance and use of technology, a model of behavior toward new technologies. It comprises 20 items, and each item was rated on a five-point scale. Using this questionnaire, we conducted an online survey in 2018 among 399 doctors and 600 members of the public. The sample-wide responses were analyzed, and then the responses of the doctors were compared with those of the public using *t* tests.

Results: Regarding the sample-wide responses (N=999), 653 (65.4%) of the respondents believed, in the future, AI in medicine would be necessary, whereas only 447 (44.7%) expressed an intention to use AI-driven medicine. Additionally, 730 (73.1%) believed that regulatory legislation was necessary, and 734 (73.5%) were concerned about where accountability lies. Regarding the comparison between doctors and the public, doctors (mean 3.43, SD 1.00) were more likely than members of the public (mean 3.23, SD 0.92) to express intention to use AI-driven medicine (P<.001), suggesting that optimism about AI in medicine is greater among doctors compared to the public.

Conclusions: Many of the respondents were optimistic about the role of AI in medicine. However, when asked whether they would like to use AI-driven medicine, they tended to give a negative response. This trend suggests that concerns about the lack of regulation and about accountability hindered acceptance. Additionally, the results revealed that doctors were more enthusiastic than members of the public regarding AI-driven medicine. For the successful implementation of AI in medicine, it would be necessary to inform the public and doctors about the relevant laws and to take measures to remove their concerns about them.

(JMIR Hum Factors 2022;9(1):e24680) doi:10.2196/24680

KEYWORDS

artificial intelligence; technology acceptance; surveys and questionnaires; doctors vs public

Introduction

Background

The use of artificial intelligence (AI) in the medical industry promises many benefits. For example, it can yield new diagnostic and therapeutic methods, provide the groundwork for introducing cutting-edge medical technology, and reduce the workload of doctors and care workers [1-3]. AI has been introduced to medicine primarily in developed countries [4]. The United States made an early start in this respect. In April 2018, the country's Food and Drug Administration (FDA) authorized the first medical device to use AI. The device, named IDx-DR, detects greater than a mild level of the eye disease diabetic retinopathy in patients with diabetes. As the FDA states, IDx-DR provides a screening decision "without the need for a clinician to also interpret the image or results" [5]. Like the United States, Japan wants to drive forward the use of AI in medicine. The country's Ministry of Health, Labour, and Welfare (MHLW) has designated six health care fields where AI is to be developed. Under the MHLW's plan, AI will be rolled out relatively early in four of these fields (genome medicine, diagnostic imaging, diagnosis and treatment, and drug development) and then in a more phased manner in the remaining two (long-term care and dementia, surgery) [1]. Despite the MHLW's efforts, however, Japan lags other developed countries in rolling out AI in health care.

As the shift toward AI in medicine continues apace, there is an urgent need to consider the ethical, legal, and social issues (ELSIs) of this trend [6]. For example, insofar as clinical data are used to develop AI applications, an issue arises regarding patients' personal data [7]. This issue has not escaped the attention of Japan; in June 2018, the MHLW released an announcement on AI-guided diagnosis, stating that AI will only ever assist a human doctor in forming a final diagnosis and that no matter how advanced AI becomes, decision-making responsibility will always lie with the human doctor [8]. Despite such reassurances, many members of the public remain concerned about where accountability lies in AI-driven medicine. Such misgivings may hinder the rollout of AI in medicine.

A new application can only fulfill its potential if people use it. Exemplifying this principle are South Korea's mobile electronic medical records (EMRs) [9]. Mobile EMRs are effective for streamlining medical work and minimizing hospital costs, but staff feel strongly disinclined to use them, and the uptake rate is low; this is because the functions are poorly tailored to the user's needs. This situation demonstrates, according to Kim [9], that an application can only fulfill its true potential if the developers consider user feedback. The success of a technology rollout depends on the technology's features, but it also depends on popular trends and the broader sociocultural milieu. Toward ensuring successful rollout, trends in public acceptance of the technology and the determinants of such would need to be identified [10,11]. When it comes to AI in medicine, there may be a gap in acceptance between the doctors, who would actually use the AI, and the public, who would receive AI-driven medical services. It is, therefore, necessary to consider acceptance among doctors and among the public. However, in Japan, the acceptance of AI in medicine has not been investigated.

Therefore, the purpose of this study was to obtain detailed data on the acceptance of AI in medicine by comparing the acceptance among Japanese doctors with that among the Japanese public.

Theoretical Background

In recent years, AI has been developed in the medical industry. For example, AI can detect pulmonary nodules, tuberculosis, and pneumonia in chest radiographs, detect and quantify pulmonary nodules in chest computed tomography (CT) [12], detect suspected large vessel occlusion strokes based on CT images [13], and screen for breast cancer [14].

The technology acceptance model (TAM) is used to investigate the acceptance of new technologies. TAM is a model that explains the process by which users accept and use information systems. There are many extended models, among which the unified theory of acceptance and use of technology (UTAUT), proposed by Venkatesh et al [15] by integrating eight models, explains 70% of the variance in individual intention to use technology while the existing technology acceptance models explain 40%. In UTAUT, the user's intention to use an information system and subsequent use is explained by four components (performance expectancy, effort expectancy, social influence, and facilitating conditions).

Literature Reviews

Oh et al [16] found that 83.4% of respondents considered that AI would be useful in the medical field, indicating that doctors have a positive attitude toward AI in medicine in a survey of doctors and medical students in Korea. Jonmarker et al [17] investigated participants' confidence in the introduction of AI into a breast cancer screening program in Sweden. Participants in a breast cancer screening program trusted the computer-aided decision-making of their doctors the most. Jutzi et al [18] also conducted a survey of patients with and without a diagnosis of melanoma to investigate the acceptance of AI for melanoma diagnostics in Germany. The results showed that only 41% agreed with the use of AI as a stand-alone system, and 94% agreed with the use of AI as a support system for doctors.

In Japan, a number of studies have polled attitudes regarding the rollout of AI in health care. In one survey by the Ministry of Internal Affairs and Communications (MIC), 81.5% of the polled experts said that they would welcome the use of AI in analyzing biometrics, lifestyle, disease history, genetic data, and other factors to detect precisely symptoms of health conditions or the onset of disease [19]. Another attitudes survey on AI in health care was conducted by Ema et al [20]. In the survey, the respondents agreed strongly with the idea of entrusting AI with driving, disaster management, military matters, and other functions where the rollout of AI requires

institutional and social consent. However, they felt that humans should remain the primary actor in matters involving individual choice, such as health management and important life decisions. As insightful as its findings are, the study had examined participants' views on the use of AI in a number of fields, not only health care. A focus on AI in medicine would present a more detailed picture of the public and expert trust in such. One study that did so was Yokoi and Nakayachi [21]; the study reported that sharing treatment plans resulted in higher trust toward AI but also that this effect was modest. However, they are investigating the reliability of AI in medicine, and there has been no investigation of acceptance focused on AI in medicine in Japan. For the successful implementation of AI in medicine in Japan, it is necessary to investigate not only the reliability but also the acceptance and factors related to acceptance.

In addition, although clarifying the differences in acceptance of AI in medicine between doctors and the public would allow us to consider approaches suitable for each of them in promoting the introduction of AI, most previous studies have focused on either doctors or the public (patients).

Our research question was to what extent AI in medicine has been accepted in Japan and whether there are differences in the acceptance of AI in medicine between doctors and the public.

Methods

Survey

For the purposes of the study, AI in medicine was defined as the use of AI to determine diagnosis and treatment without requiring a doctor.

The survey questions were divided into two sections. The first section consisted of items on the respondents' general attributes, such as sex and whether the respondent was a doctor.

The second section measured the respondents' acceptance of AI in medicine with questions referred to the UTAUT [15]. Generally, a survey is conducted for people who directly use the system, assuming a specific system in UTAUT. However, in this study, the respondents include the public who does not use the system directly. Also, we assume no specific systems because the AI in medicine was not widespread at the time the survey was conducted in Japan, and the only description of AI was "the use of AI to determine diagnosis and treatment without requiring a doctor" in the questionnaire. Therefore, we modified the questions through discussion to make them suitable for this study. Specifically, questions that are difficult to answer without assuming a specific system were deleted, and we added alternative items for four key components. In addition, since attitude and uneasiness have been verified as a factor influencing intention to use in previous studies, we added a question on attitude and uneasiness in this study [15,22]. Then, we thought it would be difficult to analyze the data by fitting the UTAUT model because we modified the questionnaire for this study, so we measured one element for all items.

There were 20 such items (Textbox 1), each representing a factor of acceptance. Each item was rated on a five-point scale (1 = completely false, 2 = somewhat false, 3 = cannot say either way, 4 = somewhat true, 5 = completely true). A question item about medical costs was rated on a different five-point scale, where 1 = costs will decrease, and 5 = costs will increase.

Textbox 1. Question items measuring acceptance factors

Usefulness: Do you think that AI in medicine will be useful?

Efficiency: If AI is used in medicine, do you think doctors could provide services more efficiently?

Better medical services: Would using AI in medicine lead to better medical services?

Mastery: Could doctors quickly master the use of AI in medicine?

User-friendliness: Could doctors easily operate AI in medicine settings?

Expectations of others: Do you think people around you are optimistic about the potential of AI in medicine?

Expectations among patients: Do you think patients are optimistic about the potential of AI in medicine?

Brand impact: Are your views on AI in medicine shaped by the businesses (or brands of such businesses) that developed the AI for medicine?

Knowledge of AI in other contexts: Do you know much about the use of AI in contexts other than medicine?

Knowledge of AI in medicine: Do you know much about the use of AI in medicine?

Medical costs: How do you think AI in medicine will affect medical costs?

Necessity of legislation: Do you think the use of AI in medicine should be regulated by legislation?

General impression: Do you have a generally favorable impression of the use of AI in medicine?

Interest in topic: Are you interested in the topic of AI in medicine?

Accuracy: Do you think AI in medicine will deliver accurate diagnoses?

Concern about data leakage: Are you concerned that using AI in medicine might lead to the leakage of personal data?

Concern about accountability: Are you concerned about who would be accountable for any accident resulting from the use of AI in medicine?

Intention to use: Would you be willing to use AI-driven medicine?

Relevance to life: Do you think that AI in medicine will play an important role in your life in the future?

Necessity in medicine: Do you think that AI will be essential in medicine in the future?

The survey was conducted over three days, from November 13 to 15, 2018. The authors did not obtain Institutional Review Board approval for this study because we used Rakuten Insight to conduct the survey, and we did not obtain any personal information from the respondents. In Japan, researchers do not have to obtain Institutional Review Board approval when subjects can voluntarily decide to participate in a study, there is no intervention in the collection of data, and individuals cannot be identified from the collected data [23]. Respondents were told the length of time to answer the questions, the purpose of the survey, and who conducted the survey on the screen just before they started to answer. Respondents were allowed to stop answering at any time until they answered all the questions. We took the completed responses as respondents' consent to participate in the survey and used the responses for analysis. In Multimedia Appendix 1, each item in the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) is shown [24].

After the questionnaire survey, the reliability of the questionnaire was examined. First, Cronbach alpha was calculated to confirm the reliability of the entire questionnaire (Cronbach α=.88). Cronbach α values of .70 to .80 or more are regarded as satisfactory [25], the questionnaire of this study was considered to be reliable.

The public and doctor responses for the 20 items were compared using a *t* test. We further calculated Cohen *d* for each item. In general, Cohen *d* of 0.2 is considered small, 0.5 is considered medium, and 0.8 is considered large [26]. In addition, the opposite scale was inverted so that a positive response to the use of AI was 5 and a negative response was 1 (eg, concern about data leakage). The total mean score of the whole scale was calculated for doctors and the public respectively and compared using a *t* test. The statistics were processed using R (version 3.5.1; R Core Team).

Respondents

An online survey was conducted among the public and among doctors. The sample representing the public survey consisted of 600 individuals across six age cohorts, each of which included 50 men and 50 women. The cohorts were aged 15 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, and 65 years and older. The sample representing doctors consisted of 400 individuals aged 25 years or older. Of these, 350 were men and 50 were women. One doctor was excluded from analysis because he answered that his highest education attainment was "vocational school/junior college," though doctors needed university education to get a license.

Regarding the respondents' general attributes, Table 1 shows the sex and age, and Table 2 shows the educational attainment of the doctors and members of the public.

Table 1. Sex and age distribution.

Age (years)	Doctors			Public		
	Male	Female	Total	Male	Female	Total
15–24	N/A[a]	N/A	N/A	50	50	100
25–34	7	13	20	50	50	100
35–44	42	19	61	50	50	100
45–54	123	11	134	50	50	100
55–64	141	5	147	50	50	100
65 or older	36	2	38	50	50	100
Total	349	50	399	300	300	600

[a]N/A: not applicable.

Table 2. Highest educational attainment.

Highest educational attainment	Doctors	Public
Junior high school	0	8
High school	0	177
Vocational school/junior college	(1)[a]	127
University	398	282
Other	1	6
Total	399	600

[a]One doctor was excluded from analysis because doctors needed university education to get a license.

Results

Table 3 shows the sample-wide results of 999 respondents for the 20 items on acceptance of AI in medicine. For the following items, under 20% of the responses were negative (“completely false” or “somewhat false”), and over 50% were positive (“completely true” or “somewhat true”): usefulness, efficiency, better medical services, expectations of others, expectations among patients, general impression, relevance to life, and necessity in medicine. The most positively rated of these items were “usefulness” and “necessity in medicine,” the number of positive responses for which were 657 (65.8%) and 653 (65.4%), respectively. One item with a relatively low positive response rate was “intention to use”; only 447 (44.7%) of the respondents indicated that they were willing to use AI-driven medicine (“I would be moderately willing to” or “I would be very willing to”).

Regarding the necessity of legislation and whether there is concern about accountability, only 360 (36.0%) and 315 (31.5%) of the respondents gave a strong affirmative response, respectively. However, when the strong and moderate responses were combined, as many as 730 (73.1%) and 734 (73.5%) gave an affirmative response, respectively.

Table 3. Sample-wide results for factors of acceptance.

Items	Mean (SD)
Usefulness	3.66 (0.91)
Efficiency	3.47 (0.89)
Better medical services	3.66 (0.97)
Mastery	3.02 (0.91)
User-friendliness	2.97 (0.91)
Expectations of others	3.44 (0.92)
Expectations among patients	3.48 (0.99)
Brand impact	3.05 (1.07)
Knowledge of AI in other contexts	2.82 (1.04)
Knowledge of AI in medicine	2.34 (1.01)
Medical costs	2.88 (0.99)
Necessity of legislation	3.99 (0.88)
General impression	3.64 (1.00)
Interest in topic	3.52 (0.81)
Accuracy	3.08 (1.10)
Concern about data leakage	3.30 (1.00)
Concern about accountability	3.92 (0.96)
Intention to use	3.31 (0.91)
Relevance to life	3.64 (0.92)
Necessity in medicine	3.70 (0.42)

Table 4 shows the comparative results for the doctors and the public. Responses for eight of the items exhibited a significant intergroup difference at the 5% level of significance: better medical services, mastery, expectations of others, knowledge of AI in medicine, medical costs, interest in the topic, concerns about data leakage, and intention to use. Intergroup difference was particularly notable in "expectations of others" and "intention to use." Among the public, the median response for both items was neutral (3. "cannot say either way)." Among doctors, the median was a moderately affirmative response (4. "I moderately agree that people around me are optimistic about AI in health care" for the former, and 4. "I would like to use AI-driven medicine in the future" for the latter). Another notable item was "knowledge of AI in medicine." The median response for this item among doctors was neutral, whereas that among the public was moderately negative (2. "I don't know all that much about it"). In addition, Cohen *d* for mastery, expectations of others, medical costs, intention to use was small, and Cohen *d* for knowledge of AI in medicine was medium.

Table 4. Comparison between doctors and the public regarding factors associated with acceptance.

Items	Doctor		Public		Cohen *d*	95% CI	*P* value
	Median	Mean (SD)	Median	Mean (SD)			
Usefulness	4	3.72 (0.93)	4	3.62 (0.90)	0.11	–0.02 to 0.21	.09
Efficiency	4	3.44 (0.89)	4	3.50 (0.86)	–0.07	–0.17 to 0.05	.30
Better medical services	4	3.73 (0.85)	4	3.61 (0.91)	0.15	0.02 to 0.24	.02
Mastery	3	3.15 (0.95)	3	2.93 (0.97)	0.23	0.10 to 0.34	<.001
User-friendliness	3	2.97 (0.91)	3	2.97 (0.90)	–0.001	–0.12 to 0.11	.98
Expectations of others	4	3.56 (0.88)	3	3.37 (0.92)	0.21	0.08 to 0.30	.001
Expectations among patients	4	3.50 (0.89)	4	3.46 (0.94)	0.05	–0.07 to 0.16	.46
Brand impact	3	3.04 (0.98)	3	3.06 (0.99)	–0.02	–0.14 to 0.11	.82
Knowledge of AI[a] in other contexts	3	2.88 (1.00)	3	2.78 (1.11)	0.10	–0.03 to 0.24	.13
Knowledge of AI in medicine	3	2.66 (1.00)	2	2.12 (1.01)	0.54	0.41 to 0.67	<.001
Medical costs	3	3.01 (0.93)	3	2.79 (1.04)	–0.23	–0.35 to –0.10	<.001
Necessity of legislation	4	4.01 (1.01)	4	3.99 (0.97)	0.02	–0.11 to 0.15	.76
General impression	4	3.69 (0.85)	4	3.60 (0.89)	0.10	–0.02 to 0.20	.13
Interest in topic	4	3.61 (1.01)	4	3.46 (0.99)	0.15	0.02 to 0.28	.02
Accuracy	3	3.11 (0.84)	3	3.06 (0.79)	0.07	–0.05 to 0.16	.32
Concern about data leakage	3	3.38 (1.08)	3	3.24 (1.11)	0.13	–0.28 to –0.002	.046
Concern about accountability	4	3.91 (1.01)	4	3.93 (0.99)	0.015	–0.11 to 0.14	.81
Intention to use	4	3.43 (1.00)	3	3.23 (0.92)	0.21	0.07 to 0.32	.002
Relevance to life	4	3.62 (0.90)	4	3.66 (0.92)	–0.04	–0.16 to 0.08	.49
Necessity in medicine	4	3.74 (0.84)	4	3.67 (0.97)	0.07	–0.05 to 0.18	.24

[a]AI: artificial intelligence.

The total mean score of the whole scale was 65.5 (SD 10.4) for doctors and 64.1 (SD 10.6) for the public, with a difference of 1.4 (95% CI 0.10-2.77). Cohen *d* was 0.14.

Discussion

Principal Results

Regarding the sample-wide results, the respondents were generally receptive toward AI in medicine. Particularly respondents have confidence in AI's usefulness and a belief in its future necessity in medicine. In the MIC survey [19], experts expressed optimism on the use of AI in diagnosis and other aspects of health care. This study revealed that optimism on AI in health care is present among the public and doctors alike, implying that such optimism is broadly held.

Despite their tendency to see AI as useful and necessary in medicine, the respondents were less enthusiastic about the prospect of actually using AI-driven medicine, with only 44.8% of the respondents giving a moderate or strong affirmative response for intention to use. According to a previous study, 41% of respondents in Germany were in favor of using AI alone to diagnose melanoma [18]. In Sweden, 38% of participants in a breast cancer screening program preferred computer-only reading [17]. Furthermore, 35.4% of Korean doctors agreed that AI could replace them in their jobs [16]. The acceptance of AI-driven medicine in Japan seems to be generally consistent with previous studies. The UTAUT model assumes that some factors encourage acceptance of technology, whereas other factors hinder such [22]. The presence of a hindering factor may be the reason that belief in AI's usefulness and necessity in medicine did not translate directly into a desire to use AI-driven medicine personally.

The majority of the sample expressed a moderate or strong concern regarding the issues of regulatory legislation and accountability. About half of the respondents expressed moderate or strong concern about data leakage. These three items describe ELSIs, which require solutions from a policy perspective. Given that so many of the respondents were concerned about these ELSIs, the ELSIs in question are likely major determinants of acceptance for both doctors and members of the public. In particular, the issue of accountability attracted concern from as many as three-fourths of the respondents, despite the MHLW's attempts to reassure people that human doctors will always be responsible for the final diagnosis. The causes of such uncertainty are unclear from this study's results; further research is necessary to identify the causes and derive ways to alleviate the concerns among doctors and the public.

Discussed below is the comparison between doctors and the public. The results revealed significant intergroup differences in eight items. One such difference was in "intention to use";

doctors were more enthusiastic than the public about using AI-driven medicine in the future. Ema et al [20] surveyed the public and 10 other stakeholders on the use of AI in eight areas, one of which was health management. In all eight areas, the study found the public to be more likely than the other stakeholders to answer that humans should remain in control. However, any comparison with this study requires some qualifications owing to salient differences. Ema et al [20] did not include doctors among the ten stakeholders and used a definition of AI different from that used in this study. Nonetheless, Ema et al's [20] observation that experts are more willing to trust AI than members of the public echoes this study's observation that doctors, compared with the public, were more receptive to the idea of using AI.

Doctors' comparative enthusiasm for using AI may be related to the fact that they were also more likely than members of the public to give positive responses to better medical services, mastery, expectations of others, and interest in the topic. That is, the doctors' intention to use AI may have been motivated by their greater expectations (compared with those held by members of the public) about the potential of AI in medicine. Additionally, members of the public were more likely than doctors to indicate a lack of knowledge about AI in medicine. The fact that members of the public tended to be rather uninformed about AI in medicine may have contributed to their weak (compared with doctors) intention to use AI.

Meanwhile, the responses to the items on medical costs and concern about data leaks present a paradox. Specifically, members of the public were more likely than doctors to believe that AI would lead to lower medical costs, whereas doctors were more likely than members of the public to express concern about the risk of data leakage. The results for these two items seem to imply that the members of the public, not the doctors, are more inclined to use AI. However, it was the doctors who gave the more affirmative responses to the actual question on the intention to use AI-driven medicine. A possible explanation for this paradox could be that the items "usefulness" and "better medical services" impact "intention to use" more than they do "medical costs" and "concern about data leakage."

Although doctors' total mean scores were significantly higher than those of the public, the effect size was negligible. This would be because a slight difference in the total mean score was detected by *t* test due to the large sample size.

Limitations

Regarding the limitations of the study, one limitation concerns the possibility of sampling bias in the online survey. Because participation in the online survey was limited to individuals who could use a personal computer, smartphone, or similar device, the sample may have been biased toward the digitally literate. Moreover, as the survey was titled "Survey on AI in medicine," the sample may have been biased toward individuals who were interested in medicine and AI. Given that people are generally more likely to express a clear opinion for or against a proposition when they are knowledgeable about the topic in question [27], a more unbiased sample may have yielded more neutral ("cannot say either way") responses. In view of these possible biases, caution is advised when interpreting the results.

Further research is necessary to explore the relations between items. This study ascertained population trends by analyzing sample-wide responses and then comparing the responses between doctors and the public. What this approach failed to clarify was the matter of which item most affects intention to use. Accordingly, future research should explore how the responses to one item correlate with those to another. In this study, we were not able to conduct an analysis using the UTAUT model. However, since AI in medicine is now starting to be used in Japan, we would like to analyze the acceptance of AI in medicine by the UTAUT model assuming a specific system in the future study.

Since this study surveyed a large sample of 399 doctors and 600 citizens, it can be considered to have at least some validity. However, it should be noted that the questionnaire was not carried out validation.

In this study, we did not investigate the health status of the public or the duration of the professional experience of doctors. In the future, it will be necessary to conduct a survey that takes this into account.

Conclusions

To the best of our knowledge, this is the first survey on the acceptance of AI in medicine in Japan. This study aimed to obtain detailed data on the acceptance of AI in medicine by comparing the acceptance among Japanese doctors with that among the Japanese public. An online survey was conducted, and the results were analyzed to determine sample-wide trends and trends specific to doctors and to the public.

In the 999 respondents, the results indicated that around two-thirds of the sample believed that AI would be useful in (657/999, 65.8%) and necessary to medicine (653/999, 65.4%). However, such beliefs did not directly translate to intention to use AI-driven medicine; only 447 (44.7%) of the sample expressed such a desire. The results also showed that 730 (73.1%) believed that regulatory legislation was necessary, and 734 (73.5%) were concerned about accountability, suggesting that these factors are important in terms of acceptance among doctors and the public alike. The comparison of the two groups revealed that doctors were more likely than members of the public to express intention to use AI-driven medicine ($P<.001$). This trend may be related to the responses for the items "better medical services," "mastery," "expectations of others," and "interest in the topic."

In this study, we did not analyze with the UTAUT model; however, the analysis with UTAUT should be done assuming a concrete system in the future.

Acknowledgments

HT, MM, and KO considered the conception and design of this research. HT and HY made an initial version of the questionnaire, and all authors revised the questionnaire. TS contributed to the acquisition of data. HT and YM analyzed the data. HT drafted the manuscript. All authors interpreted the results and revised the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Report on the Checklist for Reporting Results of Internet E-Surveys.
[PDF File (Adobe PDF File), 69 KB - humanfactors_v9i1e24680_app1.pdf]

References

1. Ministry of Health, Labour and Welfare.. AI utilization promotion meeting in the preservation of health medical field. (in Japanese). 2017. URL: https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000169230.pdf [accessed 2020-09-25]
2. Okuno Y. Data-driven drug discovery using artificial intelligence. The medical frontline. (in Japanese) 2019;74(3):374-378.
3. Seetharam K, Kagiyama N, Sengupta P. Application of mobile health, telemedicine and artificial intelligence to echocardiography. Echo Res Pract 2019 Jun 01;6(2):R41-R52 [FREE Full text] [doi: 10.1530/ERP-18-0081] [Medline: 30844756]
4. Ministry of Health, Labour and Welfare. An artificial intelligence utilization investigation report in fiscal year (the outline edition). 2017. URL: https://www.mhlw.go.jp/content/10601000/000337600.pdf [accessed 2020-09-25]
5. U.S. Food and Drug Administration. 2018. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. URL: https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye [accessed 2020-09-25]
6. Japan medical association. 2018. Artificial intelligence (AI) and medicine. URL: http://dl.med.or.jp/dl-med/teireikaiken/20180620_3.pdf [accessed 2020-09-25]
7. Vellido A. Societal Issues Concerning the Application of Artificial Intelligence in Medicine. Kidney Dis 2018 Sep 3;5(1):11-17. [doi: 10.1159/000492428]
8. Ministry of Health, Labour and Welfare. 2018. Preservation of health medical field AI development acceleration consortium. URL: https://www.mhlw.go.jp/content/10601000/000337597.pdf [accessed 2020-09-25]
9. 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]
10. Jeon YS, Kaneda R, Kano S. A study on behavioral intention for the adoption of smart phone: A survey on mobile telecommunication service users in Japan and South Korea. The Journal of Japanese Society for Global Social and Cultural Studies (in Japanese) 2010;7:27-39. [doi: 10.11424/gscs.7.1_27]
11. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A Systematic Review of the Technology Acceptance Model in Health Informatics. Appl Clin Inform 2018 Aug 15;09(03):604-634. [doi: 10.1055/s-0038-1668091]
12. Chassagnon G, Vakalopoulou M, Paragios N, Revel M. Artificial intelligence applications for thoracic imaging. Eur J Radiol 2020 Feb;123:108774. [doi: 10.1016/j.ejrad.2019.108774] [Medline: 31841881]
13. Petrone JJ. FDA approves stroke-detecting AI software. Nat Biotechnol 2018 Apr 05;36(4):290. [doi: 10.1038/nbt0418-290] [Medline: 29621226]
14. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature 2020 Jan;577(7788):89-94. [doi: 10.1038/s41586-019-1799-6] [Medline: 31894144]
15. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
16. 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]
17. Jonmarker O, Strand F, Brandberg Y, Lindholm P. The future of breast cancer screening: what do participants in a breast cancer screening program think about automation using artificial intelligence? Acta Radiol Open 2019 Dec;8(12):2058460119880315 [FREE Full text] [doi: 10.1177/2058460119880315] [Medline: 31839989]
18. Jutzi TB, Krieghoff-Henning EI, Holland-Letz T, Utikal JS, Hauschild A, Schadendorf D, et al. Artificial Intelligence in Skin Cancer Diagnostics: The Patients' Perspective. Front Med (Lausanne) 2020;7:233 [FREE Full text] [doi: 10.3389/fmed.2020.00233] [Medline: 32671078]
19. Ministry of Internal Affairs Communications. 2016. White paper information and communications in Japan. URL: http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/n4200000.pdf [accessed 2020-09-25]

20. Ema A, Akiya N, Osawa H, Hattori H, Oie S, Ichise R, et al. Future Relations between Humans and Artificial Intelligence: A Stakeholder Opinion Survey in Japan. IEEE Technol Soc Mag 2016 Dec;35(4):68-75. [doi: 10.1109/mts.2016.2618719]
21. Yokoi R, Nakayachi K. Does shared policy of medical treatment improve trust in artificial intelligence? In: The Proceedings of the Annual Convention of the JPA. 2018 Presented at: The 82nd Annual Convention of the Japanese Psychological Association; September 25, 2018 to September 27, 2018; Sendai International Center p. 1AM-021-1AM-021. [doi: 10.4992/pacjpa.82.0_1am-021]
22. Bao S, Hoshino S, Hashimoto S, Shimizu N. Factors influencing Internet adoption in rural areas of Hubei province, China. Journal of Rural Planning Association 2014;33(1):54-62. [doi: 10.2750/arp.33.54]
23. Ministry of Education, Culture, Sports, Science and Technology and Ministry of Health, Labour, and Welfare. Ethical Guidelines for Medical and Health Research Involving Human Subjects (in Japanese). 2017 Feb 28. URL: https://www.mhlw.go.jp/content/10600000/000757206.pdf
24. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Dec 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
25. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997 Feb 22;314(7080):572 [FREE Full text] [Medline: 9055718]
26. Cohen J. Statistical Power Analysis. Current Directions in Psychological Science 1992 Jun;1(3):98-101. [doi: 10.1111/1467-8721.ep10768783]
27. Kinoshita T. The structure of risk perception and its cultural difference. Journal of Japan Society for Safety Engineering (in Japanese) 2002;41(6):356-363. [doi: 10.18943/safety.41.6_356]

Abbreviations

AI: artificial intelligence
CT: computed tomography
ELSI: ethical, legal, and social issue
EMR: electronic medical record
FDA: Food and Drug Administration
MHLW: Ministry of Health, Labour, and Welfare
MIC: Ministry of Internal Affairs and Communications
TAM: technology acceptance model
UTAUT: unified theory of acceptance and use of technology

Edited by A Kushniruk; submitted 02.10.20; peer-reviewed by R Watson, J Offermann-van Heek; comments to author 02.12.20; revised version received 23.03.21; accepted 29.12.21; published 16.03.22.

Please cite as:
Tamori H, Yamashina H, Mukai M, Morii Y, Suzuki T, Ogasawara K
Acceptance of the Use of Artificial Intelligence in Medicine Among Japan's Doctors and the Public: A Questionnaire Survey
JMIR Hum Factors 2022;9(1):e24680
URL: https://humanfactors.jmir.org/2022/1/e24680
doi:10.2196/24680
PMID:35293878

Original Paper

Barriers to and Facilitators for Acceptance of Comprehensive Clinical Decision Support System–Driven Care Maps for Patients With Thoracic Trauma: Interview Study Among Health Care Providers and Nurses

Emma K Jones[1], MD; Alyssa Banks[2], MPA; Genevieve B Melton[1], MD, PhD; Carolyn M Porta[3], RN, MPH, PhD; Christopher J Tignanelli[1,4], MS, MD

[1]Department of Surgery, University of Minnesota, Minneapolis, MN, United States

[2]University of Minnesota, Minneapolis, MN, United States

[3]School of Nursing, University of Minnesota, Minneapolis, MN, United States

[4]Department of Surgery, North Memorial Health Hospital, Robbinsdale, MN, United States

Corresponding Author:
Christopher J Tignanelli, MS, MD
Department of Surgery
University of Minnesota
420 Delaware St SE
Mayo Mail Code 195
Minneapolis, MN, 55455
United States
Phone: 1 6126261968
Email: ctignane@umn.edu

Abstract

Background: Comprehensive clinical decision support (CDS) care maps can improve the delivery of care and clinical outcomes. However, they are frequently plagued by usability problems and poor user acceptance.

Objective: This study aims to characterize factors influencing successful design and use of comprehensive CDS care maps and identify themes associated with end-user acceptance of a thoracic trauma CDS care map earlier in the process than has traditionally been done. This was a planned adaptive redesign stage of a User Acceptance and System Adaptation Design development and implementation strategy for a CDS care map. This stage was based on a previously developed prototype CDS care map guided by the Unified Theory of Acceptance and Use of Technology.

Methods: A total of 22 multidisciplinary end users (physicians, advanced practice providers, and nurses) were identified and recruited using snowball sampling. Qualitative interviews were conducted, audio-recorded, and transcribed verbatim. Generation of prespecified codes and the interview guide was informed by the Unified Theory of Acceptance and Use of Technology constructs and investigative team experience. Interviews were blinded and double-coded. Thematic analysis of interview scripts was conducted and yielded descriptive themes about factors influencing the construction and potential use of an acceptable CDS care map.

Results: A total of eight dominant themes were identified: alert fatigue (theme 1), automation (theme 2), redundancy (theme 3), minimalistic design (theme 4), evidence based (theme 5), prevent errors (theme 6), comprehensive across the spectrum of disease (theme 7), and malleability (theme 8). Themes 1 to 4 addressed factors directly affecting end users, and themes 5 to 8 addressed factors affecting patient outcomes. More experienced providers prioritized a system that is *easy to use*. Nurses prioritized a system that incorporated evidence into decision support. Clinicians across specialties, roles, and ages agreed that the amount of extra work generated should be minimal and that the system should help them administer optimal care efficiently.

Conclusions: End user feedback reinforces attention toward factors that improve the acceptance and use of a CDS care map for patients with thoracic trauma. Common themes focused on system complexity, the ability of the system to fit different populations and settings, and optimal care provision. Identifying these factors early in the development and implementation process may facilitate user-centered design and improve adoption.

(JMIR Hum Factors 2022;9(1):e29019) doi:10.2196/29019

KEYWORDS

clinical decision support systems; rib fractures; trauma; Unified Theory of Acceptance and Use of Technology; human computer interaction

Introduction

Background

When designed well and implemented effectively, clinical decision support systems (CDSSs) have been shown to reduce errors in health care delivery and improve outcomes [1-3]. Clinical care maps provide disease-specific assistance to multidisciplinary clinicians to support evidence-based (EB) practice and organize care processes [4,5]. Trauma is associated with significant morbidity and mortality worldwide and variable adherence to EB practices [6-8]. In the case of rib fractures, adherence to EB practices has been shown to reduce mortality up to 3-fold [4,6]. Although trauma care maps are essential in high-volume trauma centers [2,4], most trauma patients in the United States are treated at smaller-volume community hospitals, which may be less familiar with best practices for the treatment of rib fractures [6]. Scalable clinical decision support (CDS) care maps provide an important bridge in this knowledge-to-practice gap to improve care.

Although CDSSs can be an effective tool to improve adherence to EB practices, many of these systems have posed challenges and fallen short of their full potential [6,9-12]. System challenges can most commonly be attributed to poor user design, poor implementation, or poor institutional integration, resulting in systems with low overall user acceptance [12,13]. Many CDSSs are affected by poor usability, resulting in little demonstrable improvement in adoption, in part because of a lack of integration of end-user feedback [14,15]. Improved implementation and development strategies can overcome the cited barriers that now limit acceptance. For example, the *CDS Five Rights* framework [16,17] was developed to address problems frequently encountered during CDS implementation and usage [14]. This framework ensures that planning teams focus on delivering the right information to the right person, at the right time, in the right format, and through the right channel.

The integration of user-centered design (UCD) in CDS development is another element that can improve acceptability and clinician use behavior [18]. Unfortunately, the widespread adoption of UCD is still not routine, with most UCD focusing on iterative pilot-testing during CDS development or postimplementation evaluation [13,18,19]. There is a critical need for the integration of multidisciplinary, qualitative end-user input before formal CDS electronic health record (EHR) development. Unfortunately, few such qualitative studies at this phase of CDS development exist. A recent review by Khairat et al [12] identified only 11 studies using qualitative methods to evaluate user acceptance of CDS. Only 3 of those studies evaluated acceptance during early phases, such as CDS prototyping. Furthermore, it is critical that such analyses are informed by validated theories surrounding technology intention to use behavior, such as the Technology Acceptance Model [20] or the Unified Theory of Acceptance and Use of Technology (UTAUT) [21]. To overcome these limitations, a study by Khairat et al [12] proposed the User Acceptance and System Adaptation Design (UASAD) model for CDS development, implementation, and evaluation. This model suggests acceptance can be maximized by leveraging end-user feedback (ie, quantitative via survey or qualitative) to understand and integrate user expectations and needs in early CDS development.

Despite the understanding that CDSSs improve adherence to EB practices, which in turn improve outcomes for patients with thoracic trauma, only 3 CDSSs have been published to date in thoracic trauma [2,22,23]. Of these 3 CDSSs [2], 1 CDSS is the published initial result of the implementation of the rib fracture CDS care map prototype referred to in our study and the other 2 focused neither on design nor on the implementation of the CDS.

Objectives

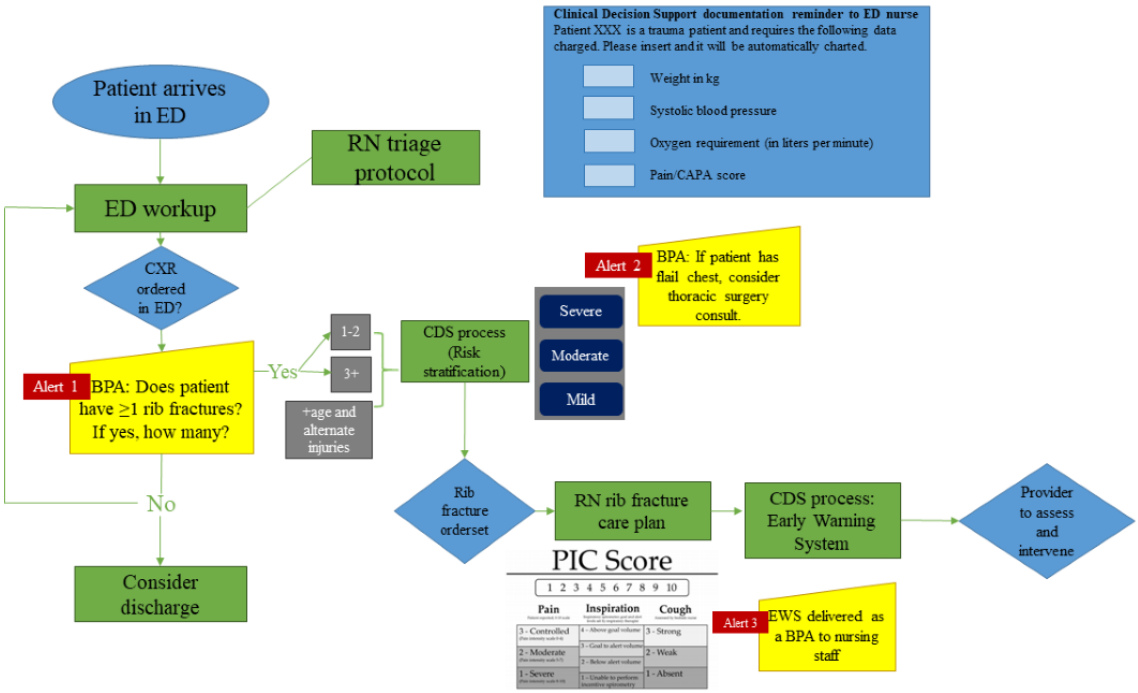
Aligned with the UASAD model for CDS development, our study uses qualitative interviews to guide the adaptive redesign of a previously created prototype CDS care map for patients with rib fractures before the formal build of the CDSS in the EHR. This is one of the earliest studies to evaluate the UASAD model for CDS prototyping and adaptive redesign. The objective of this study is to identify themes associated with end-user acceptance of a prototype CDSS guided by the UTAUT [21] model for thoracic trauma from qualitative interviews of multidisciplinary trauma end users within a 12-hospital Midwest trauma system.

Methods

Prototype CDSS Development

A multidisciplinary CDS care map planning team was assembled in April 2018 to address the care of patients with rib fractures. The planning team included 14 members with expertise in trauma surgery, trauma program management and performance improvement, trauma nursing, anesthesiology and pain management, and respiratory therapy. Over a period of 10 months, rib fracture EB practices were cataloged through a formal literature review process that identified 9 peer-reviewed articles and 1 guideline from a regional level 1 trauma center [4,24-32]. A prototype CDS care map was developed, guided by the *CDS Five Rights* framework [17] and validated EB protocols for patients with rib fractures from the American College of Surgeons Trauma Quality Improvement Project, the Eastern Association for the Surgery of Trauma, and the Western Trauma Association [5,26,33-35]. The general workflow of the prototype CDS care map is shown in Figure 1.

Figure 1. Prototype rib fracture clinical decision support (CDS) care map summary. BPA: best practice advisory; CAPA: clinically aligned pain assessment; CXR: chest x-ray; ED: emergency department; EWS: early warning system; PIC: pain inspiration cough; RN: registered nurse.



Interview Guide Development and Participants

An interview guide was constructed and included prescripted questions to identify perceptions, behaviors, barriers, and facilitators associated with both general CDS use and our proposed prototype CDSS (Multimedia Appendix 1). Interview guide development was guided by the UTAUT, a validated and widely used model that predicts the behavioral intention to use a technology and is commonly used to assess the likelihood of success for a novel technology [21]. The interview guide was tested with four end users who were not part of the final sample and revised according to their feedback: trauma inpatient registered nurse (RN), trauma advanced practice provider (APP), and 2 trauma medical doctors (MDs). A total of 22 end users comprising trauma MDs, emergency medicine MDs, APPs, and RNs were identified for participation using snowball sampling. In this paper, we use the term *providers* to refer to physicians, nurse practitioners, and physician assistants and we use the term *clinicians* to refer to all providers and nurses.

Participation was voluntary, and informed consent was obtained from participants; all interviews were audio-recorded and stored in a Health Insurance Portability and Accountability Act–compliant research environment. Before participating in the interview, the participants reviewed a 6-minute educational video describing the previously created CDS prototype. All participants completed a brief demographic survey. During the interview, the participants were asked open-ended questions guided by the interview guide. As appropriate, probing questions were asked to examine specific barriers to and facilitators for specific CDS elements.

Thematic Analysis

Interviews were transcribed verbatim using transcription software (Tybee Types Inc). A coding scheme was generated using prespecified codes based on the UTAUT constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions; emergent codes were added as appropriate. The 2 coders (AB and CJT) met weekly to develop the coding scheme and codebook. Interrater reliability was assessed through a blinded independent coding process between the 2 coders, and coding discrepancies between the coders were resolved through discussion. Following an acceptable level of agreement (>85%), all transcripts were double-coded. All transcripts were coded using computer-assisted qualitative data analysis software (NVivo). A descriptive thematic analysis approach, best described by Hsieh and Shannon [36] as conventional content analysis, was used to categorize the codes into barriers to and facilitating factors for acceptability and assess end users' intention to use the CDSS and their perspectives on the potential value of this tool.

Although interviews were semistructured, it is important to point out that specific questions regarding each of the themes were not asked; thus, the percentage of clinicians who supported each theme reflects only those clinicians who brought up a topic relating to the theme of their own volition. Therefore, the percentage values reflect the lowest possible number of clinicians, and we cannot say whether we would have had a stronger consensus had clinicians been prompted to comment on topics relating to each theme.

Ethics Approval

This study was approved by the University of Minnesota institutional review board (STUDY00005353).

Results

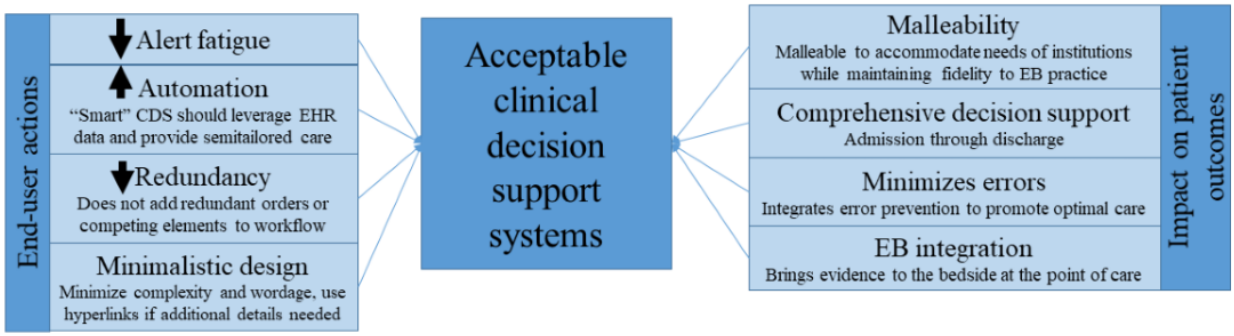
Participants

A total of 22 trauma clinicians participated, including physicians, APPs, and RNs, who were end users of the EHR. Of the 22 trauma clinicians, 11 (50%) were physicians: 3 (14%) were residents, and 8 (36%) were attending physicians, of whom all but 2 (75%) had been in practice for more than 10 years. Of the other 11 participants, 3 (27%) were APPs, of whom 1 (33%) had been practicing for more than 4 years, and 8 (73%) were RNs, of whom 6 (75%) had ≥10 years of experience. Age of the participants varied from 29 to 62 years, 73% identified as White, 9% as African American, 5% as Latinx, and 5% as multiracial. Interviewees came from a variety of practice models ranging from academic to community practice, and all but one were considered either emergency medicine or trauma surgery primary.

Thematic Analysis

The following eight major themes summarize the overarching concerns and opportunities regarding CDSSs: (1) alert fatigue, (2) automation, (3) reducing redundancy, (4) minimalistic design, (5) EB, (6) promote optimal care and prevent errors, (7) comprehensive across a spectrum of disease or injury, and (8) malleability. Each theme primarily focused on end-user actions or patient outcomes (Figure 2).

Figure 2. Eight themes for the development of acceptable clinical decision support systems (CDSS). CDS: clinical decision support; EB: evidence based; EHR: electronic health record.



Themes summarizing factors directly affecting end users include (1) alert fatigue, (2) automation, (3) reducing redundancy, and (4) minimalistic design. End users desire a *smart* experience, where orders that are already present (eg, activity, nursing, and diet orders) are not duplicated; orders are prepopulated in a risk-stratified and EB manner; and interruptive alerts or *hard stops* are only provided when absolutely necessary. Themes focused on patient outcomes include (5) EB decision support including integration of links to the evidence, practice guidelines, or web-based medical calculators; (6) promote optimal care and prevent errors; (7) comprehensive across a spectrum of disease or injury; and (8) malleability. The final theme encouraged a CDS design that allowed flexibility for each site to customize decision support and recommendations to their institutional capabilities.

We have included 1 to 2 direct quotes from qualitative interviews to illustrate each theme.

Theme 1: Alert Fatigue

The most common concern regarding CDSSs is alert fatigue. Of the 22 clinicians, 19 (86%) specifically pointed out alert fatigue as a major problem with other CDS tools they have used in practice. Their concerns surround the frequency of alerts firing, the firing of alerts at inappropriate times, and the rigidity of some alerts:

> *The other question I've got for you is: are you going to have a pop-up for rib fractures, another pop-up for spleen injury, another pop-up for femur fracture, where do you stop with the pop-ups? I'm not going to want to click through 25 pop-ups to write my admission orders.* [MD, trauma, >10 years in practice]

> *The hard-stop pop-ups are the most frustrating pop-ups because the computer does not have any judgment, so if you put a hard stop there, that computer is stopped. You can't order any more orders on your patient until you clear the pop-up. Sometimes you end up ordering something that's not appropriate for the patient just so you can continue with writing your orders.* [MD, trauma, >10 years in practice]

More experienced trauma physicians (attending physicians, 4/8, 50%) are skeptical about the added workload that CDS alerts cause and worry that they would have to find ways to work outside the system to maintain efficiency. Of the 14 providers, 6 (43%) specifically mentioned that they believe navigating pop-up alerts worsens their efficiency. Of those less concerned about alerts, 67% (2/3) primarily worked in the emergency department (ED) and consider frequent alerts helpful in their chaotic environment.

Theme 2: Automation

The second most important quality of a CDSS to end users is automation. Of the 22 clinicians, 17 (77%) believe that they should not have to search for the CDS tool; rather, it should be automatically triggered and contain prechecked orders based on EB practices. For example, medications that require renal dosing should be automatically prechecked based on the patient's glomerular filtration rate. Instead of providing providers all possible permutations of a specific disease's admission orders, the system should use risk stratification or other automated methods to precheck tailored orders for patients:

> *Boom, you can just click the box that has the dosing for the patient's weight, everything like that, so you're*

not having to go between multiple screens. [Resident MD, trauma]

RNs and providers early in their careers believe that maximally automated CDSS would improve their efficiency. Of the 11 physicians, 8 (73%) concurrently pointed out that a CDSS needs to allow for clinical judgment in fitting CDS for individual patients.

Theme 3: Redundancy

The potential for redundancy created by CDS concerns both providers and nurses; 36% (5/14) of providers and 25% (2/8) of RNs explicitly pointed out that currently used CDS tools frequently allow for overlapping orders and therefore create confusion among nurses and added work for providers who have to clean up those orders. Diet, activity, and nursing orders were frequently cited, and medication orders were less cited:

It's really important that this not generate a bunch of duplicate orders that have to be cleaned up, because that's one of the number one things that is a job dissatisfier for physicians: bogus work. [MD, trauma and critical care, >10 years in practice]

Theme 4: Minimalistic Design

Experienced providers (3/14, 21%) and some RNs (2/8, 25%) share the concern that CDS tools that are too complex (eg, poor visibility, multiple steps per task, or confusing layout) can create barriers to use. These providers value the ability to easily and quickly take in all information on a screen and have considerable disdain for EHRs that add significant complexity to their documentation experience:

I think they're cumbersome and difficult to read. They're hard to scroll through. Sometimes the scroll bar works, other times it doesn't. The words are small on the screen, and you don't see them well—they're kind of gray instead of black—so trying to read them becomes very difficult. Then there are these paragraphs and columns, and you have multiple options, and you're trying to read through these options for something, and you're scrolling through, and by the time you get to the bottom, you can't remember what the top option meant. [MD, trauma and critical care, >10 years in practice]

Theme 5: EB

Most providers (9/14, 64%) and RNs (6/8, 75%) agreed that CDS protocols must be EB. RNs specifically desired a better understanding of the evidence behind best practices:

I think there should be on the intranet something available where everybody can scroll through to see what the research is. If somebody references it, they can quick go throw in a link to it, so we can all see what it is and it's just there for everybody to see. [RN, critical care]

Physicians believe that when CDS is supported by strong evidence, it will help standardize practice across a group, particularly when it comes to less common interventions; for example, rib fixation and nerve blocks. Other perceived benefits of linking to the evidence base include increasing buy-in from users, reducing the knowledge gap between novice and expert clinicians, and helping users of different backgrounds provide consistent care:

I think that it would serve best as an advisory tool [...] here are evidence-based recommendations that will help support this patient. It takes the guesswork, especially for moonlighting physicians or providers that aren't really up on all the literature, and takes that information and moves it into the realm of recommendations that can standardize a practice across a trauma department. [MD, trauma, 4-6 years in practice]

Theme 6: Promote Optimal Care and Prevent Errors

Many clinicians (8/14, 57% of providers and 5/8, 63% of RNs) believe that CDS should help them provide optimal care and avoid errors. There is an important distinction to be made with the *EB* theme, which focuses on improving adherence to the evidence and delivering evidence at the point of care. This theme focuses on how a system can prevent errors by using alerts.

In short, the EB theme guides clinicians to what they *should do*; this theme (promote optimal care and prevent errors) prevents clinicians from doing something they *should not do*:

I think something to come and say, do you really want to do that? In order for you to move forward in this, the patient has to meet these criteria and it appears that they don't, being able to pull data from Epic to put there in front of the practitioner to say this is not in keeping with our clinical decision tool. [RN, critical care, >10 years in practice]

Institutional morbidity and mortality conferences [37] and sentinel events may provide a rich resource to guide the integration of error prevention and early warning CDSSs.

Theme 7: Comprehensive Across a Spectrum of Disease or Injury

Half of the clinicians (6/14, 43% of providers and 5/8, 63% of RNs) share the enthusiasm for a comprehensive CDSS rather than one that addresses a single-point decision. They support comprehensive disease-specific decision support from admission to discharge. Examples include incorporating tools addressing care from admission to after discharge, predicting and addressing common complications, and facilitating a multidisciplinary approach to healing:

An ideal decision support system would be both sensitive and specific, and readily identifies those patients that you may not be thinking about, and also providing you with the options of treatment that you may not be necessarily thinking about or are knowledgeable in regards to. [MD, emergency medicine, >10 years in practice]

I really think that incorporating aspects of aftercare-after hospital contact, outreach, and monitoring-would help to improve outcomes. [Resident MD, trauma]

Theme 8: Malleability

Half of the clinicians (5/14, 36% of providers and 6/8 75% of RNs) also believe that CDSSs should be malleable and able to accommodate the needs of different facilities, phases of care, and patient populations. Providers and RNs from different specialties and hospitals expressed concerns, as well as suggestions, specific to their area of practice. This was most common among ED clinicians (5/6, 83% ED clinicians). For example, our prototype CDS included an anesthesia consult for epidural placement; however, it became apparent that this would be feasible at some hospitals but not at others:

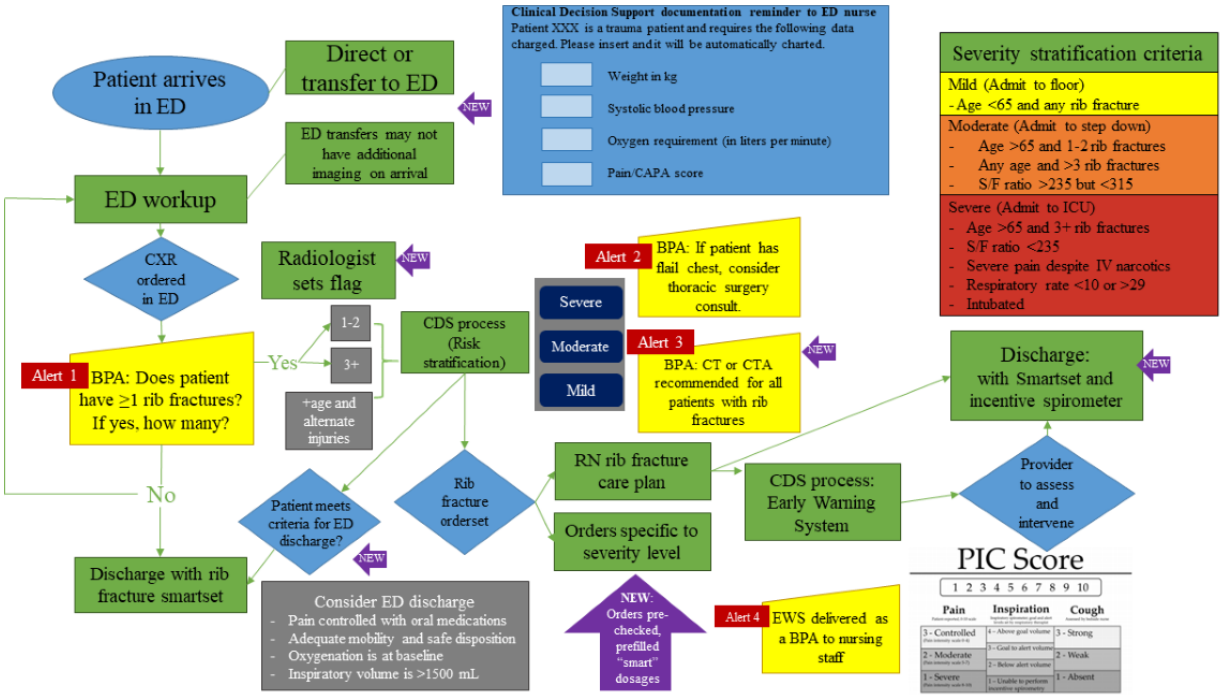
> *But we don't have a respiratory therapist, so how would that work? We don't have a pharmacist, so how would that work for us? We don't have a physical therapist or an occupational therapist, so I guess if the patient's boarding in the ER, how do we make all of that happen?* [RN, emergency medicine, >10 years in practice]

Other areas of care that were mentioned as not falling under a one-size-fits-all model were therapies, community hospitals where providers were not in-house 24/7, and older adult populations.

Select representative quotes and CDS design recommendations are given in Multimedia Appendix 2.

Using our findings from the qualitative interviews and thematic analysis, we worked with our EHR build team to implement modifications to the prototype CDS care map. For example, instead of a rib fracture multimodal analgesia panel that offers all possible analgesia options for selection *(as was initially planned in the prototype)*, end users desired a system that prepopulates analgesia based on patient risk stratification, ensures the patient does not already have duplicate medications ordered, and automatically prepopulates and prechecks the recommended nonsteroidal anti-inflammatory drug and gabapentin doses based on the patient's renal function. Figure 3 shows the workflow for the final CDSS, with new elements designated by purple arrows.

Figure 3. Final rib fracture clinical decision support (CDS) care map workflow (new elements designated by purple arrows). BPA: best practice advisory; CAPA: clinically aligned pain assessment; CDS: clinical decision support; CT: computed tomography; CTA: computed tomography angiography; CXR: chest x-ray; ED: emergency department; EWS: early warning system; PIC: pain inspiration cough; ICU: intensive care unit; IV: intravenous; RN: registered nurse; S/F: oxygen saturation/fraction of inspired oxygen ratio.

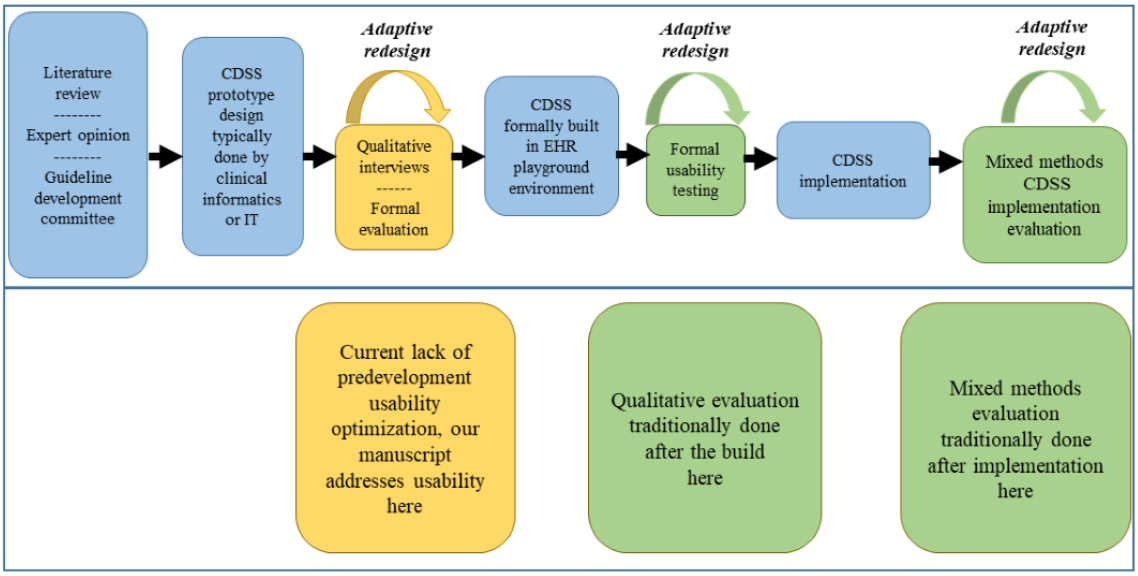


Discussion

Principal Findings

In this study, we sought to use qualitative interviews guided by the UTAUT to understand key themes associated with highly acceptable CDSSs earlier in the development and implementation process than has traditionally been done (Figure 4).

Figure 4. Process and sequence for the development of acceptable clinical decision support system (CDSS) care maps [2,3,13,38-41]. EHR: electronic health record; IT: information technology.



This was a planned stage of a UASAD development and implementation strategy focused on a 12-hospital system-wide CDS care map for patients with rib fractures. The following eight key themes were identified: alert fatigue, automation, redundancy, minimalistic design, EB, prevent errors, comprehensive CDS, and malleable design. Guided by these themes, an ideal CDS will provide support from ED arrival through discharge and not only focus on improving adherence with EB practice but also deliver evidence to the clinician at the point of care. Common errors or causes for patient decompensation that may occur when managing a specific disease process should be identified, and error prevention should be integrated into the design of the CDS. In addition, individuals responsible for the design and creation of such tools should take care to maximize automation and minimalistic design. CDSSs should leverage the rich structured EHR data available to provide more tailored and automated support; this facilitates minimalistic design as it reduces the necessity to deliver all possible orders to providers. Finally, as these systems are ultimately scaled across health care systems, it is imperative that they are designed to be easily tailored to individual hospital capabilities and resources while maintaining fidelity to EB practice.

To no surprise, alert fatigue was the most commonly cited barrier to CDSS acceptance. Our findings support the current literature that is increasingly recognizing alert fatigue as a negative consequence of CDSSs and a frequently cited reason for poor adherence [42,43]. A focused review of the current CDS literature has found no consensus on how to eliminate fatigue; however, previous studies agree on its contribution to poor adoption and clinician burnout [44-48]. Limiting the frequency of alerts or only assigning alerts to high-severity flags has been one solution previously proposed in the literature [49]. Our prototype CDSS included the following alerts that were active when the patient was in the ED: an alert prompting the radiologist to document whether rib fractures were present, an alert recommending a computed tomography scan for patients with a rib fracture seen on the chest x-ray, and an alert recommending surgery consultation for rib stabilization surgery in patients stratified to severe risk of complications. In addition, the admission order panel included 6 hard stops that forced provider action to get past.

We were able to make many recommendations to improve the final CDSS (Figure 3), which was based on the prototype shown in Figure 1 and redesigned based on our findings reported in this study. To combat alert fatigue, we recommended the integration of an artificial intelligence (AI)-enabled system that can *read* chest x-rays and tell the clinical team if the patient has rib fractures. AI diagnostic models using biomedical imaging are increasingly being investigated to improve diagnostic accuracy and minimize the workload of radiologists. They are used to facilitate imaging diagnosis for simple tasks and have been successfully used for several disease processes, including COVID-19 [50], acute respiratory distress syndrome [51], and pneumothorax [52] detection. We envision using a similar model for rib fracture detection. Using AI to perform the duties of identifying patients with rib fractures and quantifying the number of rib fractures present could remove these tasks from providers and decrease the frequency of alerts. To address the computed tomography scan and surgery consultation alert recommendations, we proposed a solution that monitors provider adherence with specific EB practices. In our final CDSS, clinicians that are already adherent to EB practice over a prespecified threshold will cease to receive notifications unless adherence falls below the threshold. Finally, all 6 hard stops were removed from the admission order panel. Interestingly, most of the clinicians who had expressed less concern with alert fatigue worked primarily in the ED and believed frequent alerts to be a generally positive thing and helpful in their chaotic environment. This suggests that different specialties may have different thresholds for tolerating interruptive decision support and needs further investigation.

Similarly, clinicians from all backgrounds agreed that the ideal CDSS should be maximally automated. These findings support the current literature that has shown that clinicians are hesitant

to use CDSSs that require additional time and effort [53]. Another study found that automated decision support was 1 of 4 main factors contributing to the success of CDS [54]. The strongest enthusiasm for this came from younger providers who referenced their workload and need for the EHR to improve rather than hinder their efficiency. However, experienced users were more likely to mention the necessity of clinical judgment. The necessity of provider judgment in high-acuity situations has been previously identified as a challenge when designing CDS, and it is imperative to build flexibility into CDS to allow for clinician judgment [55]. Interestingly, experienced (>10 years in practice) providers were more likely to have a negative association with tools that were too complex or lengthy in appearance, suggesting that as new physicians move into practice, there will be a shift toward increased tolerance of technology complexity. In today's EHR, when entering orders, providers frequently use *order sets* that include selectable orders related to the patient's disease process. Historically, order sets contain lists of orders for the provider to select from; the orders are not prechecked, and thus order sets require significant time on the part of providers to decide the orders they want and then to select each individual order. By prepopulating and prechecking orders, we are providing cognitive support so providers can easily see which orders are recommended for a disease process, thereby minimizing their workload and time spent checking boxes. Although the prototype CDSS already included an automated machine risk stratification (ie, mild, moderate, and severe) that presented individualized admission order panels, in the final CDSS, we further supplemented this by automating medication dosing, order set integration, and the calculation of a rib fracture decompensation scoring system.

Historically, CDSSs have addressed individual components of a patient's care (eg, an admission order set or ordering a specific test), which can result in disjointed care. Most CDS developed to date focus on single decision points and less on comprehensive disease-specific decision support spanning the duration of hospitalization. As a solution, we suggest creating a multifaceted CDSS that addresses care across the spectrum of a disease; this has also been suggested in the literature but has not been extensively studied [56,57]. A CDSS that incorporates all phases of care, from admission orders and imaging to discharge and follow-up, could be built to avoid redundancy as well. In addition, a CDSS with components that target different members of a multidisciplinary team may improve processes in today's increasingly team-centered health care model [48]. To this effect, our final CDSS, which was modified based on our findings reported here, includes decision support modules for ED providers, ED RNs, respiratory therapists, the admitting team (ie, trauma surgery or internal medicine), and inpatient RNs. For ED nursing, ED decision support centers on a collection of elements critical for risk stratification, whereas inpatient nursing leverages the Epic EHR *Nurse Brain*. Decision support for inpatient providers centers on support for admission, detection of clinical worsening, and discharge; for ED providers, it assists in identifying patients with rib fractures and triaging them appropriately.

To address redundancy, our final CDSS is not delivered as another admission order set, but rather as an integrated order panel within standard admission order sets. To promote EB care, links are included in all EB guidelines or decision aids when relevant. To reduce errors and promote optimal care, medication alerts were created to trigger in response to abnormally high dosages, and medications that require renal dosing will cross-reference the patient's glomerular filtration rate. To promote optimal care, an early warning system specifically tailored for patients with rib fractures [4,58] is integrated into the CDS to identify patients at high risk for decompensation and prompt early intervention. Finally, although the prototype CDSS was originally developed for a tertiary academic trauma center, the final CDSS was subsequently tailored to maintain fidelity but optimize logistics for various Midwest community hospitals. For example, in lieu of neuraxial blockade, regional catheters can be more easily provided at certain sites, and in lieu of the intensive care unit admission for all older adult patients with multiple rib fractures, a dedicated respiratory unit or specialized trauma floor may be used.

The integration of formal end-user qualitative feedback guided by the UTAUT model resulted in significant redesign of a trauma CDS care map and provided a framework for future prototyping. While pilot-testing or usability testing, a formal process of observing end-user interactions with a system to identify problems to repair and measure user performance [59], is an integral part of a UCD for CDS; most studies focus on UCD *after* the EHR build is complete with iterative redesign before implementation [12,18]. We believe that using qualitative assessment as a component of the UASAD model before the EHR build improves UCD by engaging end users early in the process. These recommendations are especially timely, as the realm of decision support for patients with thoracic trauma has not yet been extensively explored, with few existing CDSs reported [2,22,23].

Conclusions

In this study, we identified the benefit of using the UASAD model in the development of an acceptable prototype CDS care map before the EHR build and formal usability testing. By optimizing user acceptance through this qualitative method of prototype design before the EHR build, the UASAD model may result in fewer iterative redesigns during usability testing and ultimately reduce development time. In addition, this approach can accommodate the input of many multidisciplinary end users, facilitating the generalizability of user acceptance. Finally, it is possible that the integration of a UTAUT-driven qualitative redesign may facilitate more substantive CDS modifications, as usability testing typically focuses on optimizing how users complete certain tasks or interact with the CDS. UASAD-adaptive CDS redesign may offer end users a blank slate to maximize acceptance and tailor initial EHR build to institutional resources and workflow. Our experience with qualitative assessment of a prototype CDS care map has helped us identify 8 themes associated with acceptable CDS that may be used as a framework for future CDS design. CDS adaptive redesign guided by end-user qualitative analysis and validated technology acceptance theories may result in systems with higher acceptability. Further research is needed to identify specific ways to incorporate these features into CDS and evaluate trends in outcomes. Our team's next steps in the

development and implementation of a rib fracture CDS care map involve performing formal usability testing on the final CDSS, iterative redesign based on findings, implementation, and assessment of outcomes.

Limitations

Our study focused specifically on an inpatient CDS care map; therefore, these recommendations may not be generalizable for the ambulatory setting. We limited our clinicians and prototype to the trauma population; therefore, the findings may not be generalizable to nontrauma patients. Although interviewees covered both university and community hospital settings, rural and federal hospitals were not well-represented in the sample; thus, the findings may not be applicable to those settings.

Acknowledgments

The authors would like to thank Fairview Health Services IT staff, the M Health Fairview Optimal Care Program, and the University of Minnesota Center for Learning Health System Sciences.

This research was supported by the Agency for Healthcare Research and Quality and Patient-Centered Outcomes Research Institute under grant K12HS026379 (CJT), Agency for Healthcare Research and Quality grant R01HS24532 (GBM), and the University of Minnesota Clinical and Translational Science Award grant UL1TR000114.

Authors' Contributions

EKJ was involved with study design, data analysis, data interpretation, writing, and critical revision. AB was involved in study design, data collection, data analysis, data interpretation, and critical revision. GBM played a role in study design, data interpretation, and critical revision. CMP was involved study design, data interpretation, and critical revision. CJT was involved study design, data collection, data analysis, data interpretation, writing, and critical revision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide and script.
[DOCX File , 28 KB - humanfactors_v9i1e29019_app1.docx]

Multimedia Appendix 2
Representative quotes by theme.
[DOCX File , 21 KB - humanfactors_v9i1e29019_app2.docx]

References

1. Burack RC, Gimotty PA, George J, Simon MS, Dews P, Moncrease A. The effect of patient and physician reminders on use of screening mammography in a health maintenance organization. Results of a randomized controlled trial. Cancer 1996 Oct 15;78(8):1708-1721. [Medline: 8859184]
2. Macheel C, Reicks P, Sybrant C, Evans C, Farhat J, West MA, et al. Clinical decision support intervention for rib fracture treatment. J Am Coll Surg 2020 Aug;231(2):249-256. [doi: 10.1016/j.jamcollsurg.2020.04.023] [Medline: 32360959]
3. Nguyen AS, Yang S, Thielen BV, Techar K, Lorenzo RM, Berg C, et al. Clinical decision support intervention and time to imaging in older patients with traumatic brain injury. J Am Coll Surg 2020 Sep;231(3):361-367. [doi: 10.1016/j.jamcollsurg.2020.05.023] [Medline: 32561447]
4. Todd SR, McNally MM, Holcomb JB, Kozar RA, Kao LS, Gonzalez EA, et al. A multidisciplinary clinical pathway decreases rib fracture-associated infectious morbidity and mortality in high-risk trauma patients. Am J Surg 2006 Dec;192(6):806-811. [doi: 10.1016/j.amjsurg.2006.08.048] [Medline: 17161098]
5. Brasel KJ, Moore EE, Albrecht RA, deMoya M, Schreiber M, Karmy-Jones R, et al. Western trauma association critical decisions in trauma: management of rib fractures. J Trauma Acute Care Surg 2017 Jan;82(1):200-203. [doi: 10.1097/TA.0000000000001301] [Medline: 27779590]
6. Tignanelli CJ, Rix A, Napolitano LM, Hemmila MR, Ma S, Kummerfeld E. Association between adherence to evidence-based practices for treatment of patients with traumatic rib fractures and mortality rates among US trauma centers. JAMA Netw Open 2020 Mar 02;3(3):e201316 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.1316] [Medline: 32215632]
7. Tignanelli CJ, Kolk WE, Mikhail JN, Delano MJ, Hemmila MR. Noncompliance with American College of Surgeons Committee on Trauma recommended criteria for full trauma team activation is associated with undertriage deaths. J Trauma Acute Care Surg 2018 Feb;84(2):287-294. [doi: 10.1097/TA.0000000000001745] [Medline: 29360717]

8. Oliphant BW, Tignanelli CJ, Napolitano LM, Goulet JA, Hemmila MR. American College of Surgeons Committee on Trauma verification level affects trauma center management of pelvic ring injuries and patient mortality. J Trauma Acute Care Surg 2019 Jan;86(1):1-10. [doi: 10.1097/TA.0000000000002062] [Medline: 30188423]
9. Demakis JG, Beauchamp C, Cull WL, Denwood R, Eisen SA, Lofgren R, et al. Improving residents' compliance with standards of ambulatory care: results from the VA Cooperative Study on Computerized Reminders. J Am Med Assoc 2000 Sep 20;284(11):1411-1416. [doi: 10.1001/jama.284.11.1411] [Medline: 10989404]
10. Dexter PR, Perkins S, Overhage JM, Maharry K, Kohler RB, McDonald CJ. A computerized reminder system to increase the use of preventive care for hospitalized patients. N Engl J Med 2001 Sep 27;345(13):965-970. [doi: 10.1056/NEJMsa010181] [Medline: 11575289]
11. Feldstein A, Elmer PJ, Smith DH, Herson M, Orwoll E, Chen C, et al. Electronic medical record reminder improves osteoporosis management after a fracture: a randomized, controlled trial. J Am Geriatr Soc 2006 Mar;54(3):450-457. [doi: 10.1111/j.1532-5415.2005.00618.x] [Medline: 16551312]
12. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
13. Slight SP, Seger DL, Nanji KC, Cho I, Maniam N, Dykes PC, et al. Are we heeding the warning signs? Examining providers' overrides of computerized drug-drug interaction alerts in primary care. PLoS One 2013;8(12):e85071 [FREE Full text] [doi: 10.1371/journal.pone.0085071] [Medline: 24386447]
14. Liberati EG, Ruggiero F, Galuppo L, Gorli M, González-Lorenzo M, Maraldi M, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. Implement Sci 2017 Sep 15;12(1):113 [FREE Full text] [doi: 10.1186/s13012-017-0644-2] [Medline: 28915822]
15. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, et al. Effect of clinical decision-support systems: a systematic review. Ann Intern Med 2012 Jul 3;157(1):29-43. [doi: 10.7326/0003-4819-157-1-201207030-00450] [Medline: 22751758]
16. Sirajuddin AM, Osheroff JA, Sittig DF, Chuo J, Velasco F, Collins DA. Implementation pearls from a new guidebook on improving medication use and outcomes with clinical decision support. Effective CDS is essential for addressing healthcare performance improvement imperatives. J Healthc Inf Manag 2009;23(4):38-45 [FREE Full text] [Medline: 19894486]
17. Osheroff J, Teich J, Levick D, Saldana L, Velasco F, Sittig D, et al. Improving Outcomes With Clinical Decision Support: An Implementer's Guide. Boca Raton: IMSS Publishing; 2012.
18. Brunner J, Chuang E, Goldzweig C, Cain CL, Sugar C, Yano EM. User-centered design to improve clinical decision support in primary care. Int J Med Inform 2017 Dec;104:56-64 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.05.004] [Medline: 28599817]
19. Ellsworth MA, Dziadzko M, O'Horo JC, Farrell AM, Zhang J, Herasevich V. An appraisal of published usability evaluations of electronic health records via systematic review. J Am Med Inform Assoc 2017 Jan;24(1):218-226. [doi: 10.1093/jamia/ocw046] [Medline: 27107451]
20. 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]
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. Choi J, Marafino BJ, Vendrow EB, Tennakoon L, Baiocchi M, Spain DA, et al. Rib fracture frailty index: a risk-stratification tool for geriatric patients with multiple rib fractures. J Trauma Acute Care Surg 2021 Aug 25;91(6):932-939. [doi: 10.1097/TA.0000000000003390] [Medline: 34446653]
23. Kim J, Kim S, Kim YJ, Kim KG, Park J. Quantitative measurement method for possible rib fractures in chest radiographs. Healthc Inform Res 2013 Sep;19(3):196-204 [FREE Full text] [doi: 10.4258/hir.2013.19.3.196] [Medline: 24175118]
24. Bulger EM, Arneson MA, Mock CN, Jurkovich GJ. Rib fractures in the elderly. J Trauma 2000 Jun;48(6):1040-1046. [doi: 10.1097/00005373-200006000-00007] [Medline: 10866248]
25. Carver TW, Milia DJ, Somberg C, Brasel K, Paul J. Vital capacity helps predict pulmonary complications after rib fractures. J Trauma Acute Care Surg 2015 Sep;79(3):413-416. [doi: 10.1097/TA.0000000000000744] [Medline: 26307874]
26. Galvagno SM, Smith CE, Varon AJ, Hasenboehler EA, Sultan S, Shaefer G, et al. Pain management for blunt thoracic trauma: a joint practice management guideline from the Eastern Association for the Surgery of Trauma and Trauma Anesthesiology Society. J Trauma Acute Care Surg 2016 Nov;81(5):936-951. [doi: 10.1097/TA.0000000000001209] [Medline: 27533913]
27. McGonigal M. Rib fracture management. The Trauma Professional's Blog. 2012. URL: http://regionstraumapro.com/post/7843371333 [accessed 2017-06-02]
28. Pieracci FM, Coleman J, Ali-Osman F, Mangram A, Majercik S, White TW, et al. A multicenter evaluation of the optimal timing of surgical stabilization of rib fractures. J Trauma Acute Care Surg 2018 Jan;84(1):1-10. [doi: 10.1097/TA.0000000000001729] [Medline: 29077677]
29. Rieg AD, Stoppe C, Rossaint R, Coburn M, Hein M, Schälte G. [EzPAP® therapy of postoperative hypoxemia in the recovery room : experiences with the new compact system of end-expiratory positive airway pressure]. Anaesthesist 2012 Oct;61(10):867-874. [doi: 10.1007/s00101-012-2083-4] [Medline: 23011043]

30. Simon BJ, Cushman J, Barraco R, Lane V, Luchette FA, Miglietta M, EAST Practice Management Guidelines Work Group. Pain management guidelines for blunt thoracic trauma. J Trauma 2005 Nov;59(5):1256-1267. [doi: 10.1097/01.ta.0000178063.77946.f5] [Medline: 16385313]
31. Simon B, Ebert J, Bokhari F, Capella J, Emhoff T, Hayward T, Eastern Association for the Surgery of Trauma. Management of pulmonary contusion and flail chest: an Eastern Association for the Surgery of Trauma practice management guideline. J Trauma Acute Care Surg 2012 Nov;73(5 Suppl 4):351-361. [doi: 10.1097/TA.0b013e31827019fd] [Medline: 23114493]
32. Warner R, Cornell N, Knollinger P, Wilson A. Is an FVC of 1.5 adequate for predicting respiratory sufficiency in rib fractures? Scientific Session I - West Virginia University. 2017. URL: https://www.east.org/content/documents/2017_scientificpapers_webfile.pdf [accessed 2022-02-24]
33. Kasotakis G, Hasenboehler EA, Streib EW, Patel N, Patel MB, Alarcon L, et al. Operative fixation of rib fractures after blunt trauma: a practice management guideline from the Eastern Association for the Surgery of Trauma. J Trauma Acute Care Surg 2017 Mar;82(3):618-626. [doi: 10.1097/TA.0000000000001350] [Medline: 28030502]
34. Trauma quality improvement project. American College of Surgeons. 2020. URL: https://www.facs.org/quality-programs/trauma/tqp/center-programs/tqip [accessed 2022-02-24]
35. Best practices guidelines in imaging. American College of Surgeons, Trauma Quality Improvement Project. 2020. URL: https://www.facs.org/-/media/files/quality-programs/trauma/tqip/imaging_guidelines.ashx [accessed 2022-02-24]
36. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
37. Tignanelli CJ, Embree GG, Barzin A. House staff-led interdisciplinary morbidity and mortality conference promotes systematic improvement. J Surg Res 2017 Jun 15;214:124-130. [doi: 10.1016/j.jss.2017.02.065] [Medline: 28624033]
38. Heselmans A, Aertgeerts B, Donceel P, Geens S, Van de Velde S, Ramaekers D. Family physicians' perceptions and use of electronic clinical decision support during the first year of implementation. J Med Syst 2012 Dec;36(6):3677-3684. [doi: 10.1007/s10916-012-9841-3] [Medline: 22402980]
39. Kortteisto T, Komulainen J, Mäkelä M, Kunnamo I, Kaila M. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. BMC Health Serv Res 2012 Oct 08;12:349 [FREE Full text] [doi: 10.1186/1472-6963-12-349] [Medline: 23039113]
40. Patterson ES, Doebbeling BN, Fung CH, Militello L, Anders S, Asch SM. Identifying barriers to the effective use of clinical reminders: bootstrapping multiple methods. J Biomed Inform 2005 Jun;38(3):189-199 [FREE Full text] [doi: 10.1016/j.jbi.2004.11.015] [Medline: 15896692]
41. Pope C, Halford S, Turnbull J, Prichard J, Calestani M, May C. Using computer decision support systems in NHS emergency and urgent care: ethnographic study using normalisation process theory. BMC Health Serv Res 2013 Mar 23;13:111 [FREE Full text] [doi: 10.1186/1472-6963-13-111] [Medline: 23522021]
42. 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 [FREE Full text] [doi: 10.1186/s12911-017-0430-8] [Medline: 28395667]
43. 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]
44. Blecker S, Pandya R, Stork S, Mann D, Kuperman G, Shelley D, et al. Interruptive versus noninterruptive clinical decision support: usability study. JMIR Hum Factors 2019 Apr 17;6(2):e12469 [FREE Full text] [doi: 10.2196/12469] [Medline: 30994460]
45. Long D, Capan M, Mascioli S, Weldon D, Arnold R, Miller K. Evaluation of user-interface alert displays for clinical decision support systems for sepsis. Crit Care Nurse 2018 Aug;38(4):46-54 [FREE Full text] [doi: 10.4037/ccn2018352] [Medline: 30068720]
46. Jankovic I, Chen JH. Clinical decision support and implications for the clinician burnout crisis. Yearb Med Inform 2020 Aug;29(1):145-154 [FREE Full text] [doi: 10.1055/s-0040-1701986] [Medline: 32823308]
47. Chaparro JD, Hussain C, Lee JA, Hehmeyer J, Nguyen M, Hoffman J. Reducing interruptive alert burden using quality improvement methodology. Appl Clin Inform 2020 Jan;11(1):46-58. [doi: 10.1055/s-0039-3402757] [Medline: 31940671]
48. Carayon P, Hoonakker P. Human factors and usability for health information technology: old and new challenges. Yearb Med Inform 2019 Aug;28(1):71-77 [FREE Full text] [doi: 10.1055/s-0039-1677907] [Medline: 31419818]
49. Horsky J, Phansalkar S, Desai A, Bell D, Middleton B. Design of decision support interventions for medication prescribing. Int J Med Inform 2013 Jun;82(6):492-503. [doi: 10.1016/j.ijmedinf.2013.02.003] [Medline: 23490305]
50. Sun J, Peng L, Li T, Adila D, Zaiman Z, Melton G, et al. A prospective observational study to investigate performance of a chest x-ray artificial intelligence diagnostic support tool across 12 U.S. hospitals. medRxiv. 2021. URL: https://www.medrxiv.org/content/10.1101/2021.06.04.21258316v1 [accessed 2022-02-24]
51. Sjoding MW, Taylor D, Motyka J, Lee E, Co I, Claar D, et al. Deep learning to detect acute respiratory distress syndrome on chest radiographs: a retrospective study with external validation. Lancet Digit Health 2021 Jun;3(6):340-348 [FREE Full text] [doi: 10.1016/S2589-7500(21)00056-X] [Medline: 33893070]

52. Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. PLoS Med 2018 Nov;15(11):e1002697 [FREE Full text] [doi: 10.1371/journal.pmed.1002697] [Medline: 30457991]
53. Aronsky D, Haug PJ. An integrated decision support system for diagnosing and managing patients with community-acquired pneumonia. Proc AMIA Symp 1999:197-201 [FREE Full text] [Medline: 10566348]
54. 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. Br Med J 2005 Apr 2;330(7494):765 [FREE Full text] [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]
55. Williams CN, Bratton SL, Hirshberg EL. Computerized decision support in adult and pediatric critical care. World J Crit Care Med 2013 Nov 04;2(4):21-28 [FREE Full text] [doi: 10.5492/wjccm.v2.i4.21] [Medline: 24701413]
56. Austin SM, Balas EA, Mitchell JA, Ewigman BG. Effect of physician reminders on preventive care: meta-analysis of randomized clinical trials. Proc Annu Symp Comput Appl Med Care 1994:121-124 [FREE Full text] [Medline: 7949904]
57. Saperstein Y, Ong SY, Al-Bermani T, Park J, Saperstein Y, Olayinka J, et al. COVID-19 guidelines changing faster than the virus: implications of a clinical decision support app. Int J Clin Res Trials 2020;5(2):148 [FREE Full text] [doi: 10.15344/2456-8007/2020/148] [Medline: 32832740]
58. Witt CE, Bulger EM. Comprehensive approach to the management of the patient with multiple rib fractures: a review and introduction of a bundled rib fracture management protocol. Trauma Surg Acute Care Open 2017;2(1):e000064 [FREE Full text] [doi: 10.1136/tsaco-2016-000064] [Medline: 29766081]
59. Schumacher R, Lowry S. NIST guide to the processes approach for improving the usability of electronic health records. US Department of Commerce, National Institute of Standards and Technology. 2010. URL: https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf [accessed 2022-02-22]

Abbreviations

AI: artificial intelligence
APP: advance practice provider
CAPA: clinically aligned pain assessment
CDS: clinical decision support
CDSS: clinical decision support system
EB: evidence based
ED: emergency department
EHR: electronic health record
MD: medical doctor
RN: registered nurse
UASAD: User Acceptance and System Adaptation Design
UCD: user-centered design
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by A Kushniruk; submitted 22.03.21; peer-reviewed by D Chrimes, A Joseph, E Stringer; comments to author 02.10.21; revised version received 04.11.21; accepted 19.12.21; published 16.03.22.

Please cite as:
Jones EK, Banks A, Melton GB, Porta CM, Tignanelli CJ
Barriers to and Facilitators for Acceptance of Comprehensive Clinical Decision Support System–Driven Care Maps for Patients With Thoracic Trauma: Interview Study Among Health Care Providers and Nurses
JMIR Hum Factors 2022;9(1):e29019
URL: https://humanfactors.jmir.org/2022/1/e29019
doi:10.2196/29019
PMID:35293873

Original Paper

A Patient Outcomes–Driven Feedback Platform for Emergency Medicine Clinicians: Human-Centered Design and Usability Evaluation of Linking Outcomes Of Patients (LOOP)

Alexandra T Strauss[1], MIE, MD; Cameron Morgan[2], MA; Christopher El Khuri[3], MD; Becky Slogeris[2], MA; Aria G Smith[3], MS; Eili Klein[3], PhD; Matt Toerper[3,4], BSc; Anthony DeAngelo[4], MHA; Arnaud Debraine[4], BS; Susan Peterson[3], MD; Ayse P Gurses[3,5,6], MS, MPH, PhD; Scott Levin[3,4], PhD; Jeremiah Hinson[3,4], MD, PhD

[1]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[2]Center for Social Design, Maryland Institute College of Art, Baltimore, MD, United States

[3]Department of Emergency Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States

[4]StoCastic, Towson, MD, United States

[5]Armstrong Institute Center for Health Care Human Factors, Johns Hopkins Medicine, Baltimore, MD, United States

[6]Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Corresponding Author:
Alexandra T Strauss, MIE, MD
Department of Medicine
Johns Hopkins University School of Medicine
600 N Wolfe St
Block 465
Baltimore, MD, 21287
United States
Phone: 1 6098280943
Email: atstrauss13@gmail.com

Abstract

Background: The availability of patient outcomes–based feedback is limited in episodic care environments such as the emergency department. Emergency medicine (EM) clinicians set care trajectories for a majority of hospitalized patients and provide definitive care to an even larger number of those discharged into the community. EM clinicians are often unaware of the short- and long-term health outcomes of patients and how their actions may have contributed. Despite large volumes of patients and data, outcomes-driven learning that targets individual clinician experiences is meager. Integrated electronic health record (EHR) systems provide opportunity, but they do not have readily available functionality intended for outcomes-based learning.

Objective: This study sought to unlock insights from routinely collected EHR data through the development of an individualizable patient outcomes feedback platform for EM clinicians. Here, we describe the iterative development of this platform, Linking Outcomes Of Patients (LOOP), under a human-centered design framework, including structured feedback obtained from its use.

Methods: This multimodal study consisting of human-centered design studios, surveys (24 physicians), interviews (11 physicians), and a LOOP application usability evaluation (12 EM physicians for ≥30 minutes each) was performed between August 2019 and February 2021. The study spanned 3 phases: (1) conceptual development under a human-centered design framework, (2) LOOP technical platform development, and (3) usability evaluation comparing pre- and post-LOOP feedback gathering practices in the EHR.

Results: An initial human-centered design studio and EM clinician surveys revealed common themes of disconnect between EM clinicians and their patients after the encounter. Fundamental postencounter outcomes of death (15/24, 63% respondents identified as useful), escalation of care (20/24, 83%), and return to ED (16/24, 67%) were determined high yield for demonstrating proof-of-concept in our LOOP application. The studio aided the design and development of LOOP, which integrated physicians throughout the design and content iteration. A final LOOP prototype enabled usability evaluation and iterative refinement prior to launch. Usability evaluation compared to status quo (ie, pre-LOOP) feedback gathering practices demonstrated a shift across all outcomes from “not easy” to “very easy” to obtain and from “not confident” to “very confident” in estimating outcomes after using LOOP. On a scale from 0 (unlikely) to 10 (most likely), the users were very likely (9.5) to recommend LOOP to a colleague.

Conclusions: This study demonstrates the potential for human-centered design of a patient outcomes–driven feedback platform for individual EM providers. We have outlined a framework for working alongside clinicians with a multidisciplined team to develop and test a tool that augments their clinical experience and enables closed-loop learning.

(JMIR Hum Factors 2022;9(1):e30130) doi:10.2196/30130

KEYWORDS

emergency medicine; usability; human-centered design; health informatics; feedback; practice-based learning and improvement; emergency room; ER; platform; outcomes; closed-loop learning

Introduction

Proficiency in the practice of medicine is achieved over years of rigorous training and is maintained through a lifelong commitment to practice-based learning and improvement [1]. This is among the core competencies described by the Accreditation Council for Graduate Medical Education (ACGME) for all physician trainees and has been integrated into the Maintenance of Certification program by the American Board of Medical Specialties (ABMS) [2,3]. To engage in practice-based learning, clinicians must continuously assess the effectiveness of their own clinical practice [4] and actively work to make improvements at the individual and system levels [5]. Optimal experiential learning requires robust feedback mechanisms; therefore, the learner understands the real-world consequences of the actions taken and is provided the opportunity to correct course in response to suboptimal outcomes [6]. This type of closed-loop learning is a core component of deliberate practice and has been central to medical education since the time of William Osler [7]. However, the availability of outcomes-based feedback is highly variable across practice settings and medical specialties [8,9] and clinicians are often unaware of how their actions affect the short- and long-term health of patients [10].

Emergency medicine (EM) clinicians play a pivotal role in the health care system; yet, practice in an environment makes outcomes-driven learning particularly challenging. Emergency departments (EDs) are a point of entry for acutely injured and critically ill patients and are a primary source of health care for vulnerable populations [11]. In 2016, ED encounters exceeded 145 million in the United States alone [12]. While making high-stakes decisions under excessive cognitive loading and time pressure [12-15], EM clinicians set care trajectories for the majority of hospitalized patients and provide definitive care to an even larger number who are discharged into the community [16]. Additionally, because of the episodic nature of emergency care, longitudinal doctor-patient relationships do not exist in the ED. Currently, there is no mechanism for delivering systematic information about post-ED encounter patient outcomes to emergency clinicians for patient outcome–based feedback [10,17]. Emergency clinicians recognize the value of practice-based and outcomes-informed experiential learning, and they are interested in more robust postencounter feedback systems. These systems have shown potential to decrease adverse ED events, improve team function, and further clinician professional development [3,6,17,18]. Currently, postencounter telephone calls to patients of interest and case conferences (eg, morbidity, mortality) are the most common methods used to elicit postencounter patient outcome feedback in EM [17]. Over the past decade, the widespread adoption of electronic health records (EHRs) has generated continuously growing pools of clinical data, including data related to post-ED encounter patient outcomes with the potential to inform clinician practice and facilitate practice-based learning [19,20]. To date, this potential has not been realized.

In this study, we sought to unlock insights from routinely collected EHR data through the development of an individualizable patient outcomes–feedback platform. Here, we describe the iterative development of this platform, Linking Outcomes of Patients (LOOP), under a human-centered design (HCD) framework [21] and execute this through a unique collaboration between the Johns Hopkins Schools of Medicine and Engineering as well as Maryland Institute College of Art (MICA). We also report the functionality and usability of LOOP as assessed by the direct measurement of the end user clinicians' knowledge, skills, and attitudes as they interacted with and used LOOP.

Methods

Research Team Structure and Study Population

This mixed methods study was performed between August 2019 and February 2021 via a collaborative effort between the Center for Social Design at MICA and the Center for Data Science in Emergency Medicine (CDEM) at the Johns Hopkins University School of Medicine in Baltimore, Maryland. Our core study team comprised EM physicians, design researchers, human factors engineers, software engineers, and data analysts. This project was conducted in 3 phases: (1) conceptual development under an HCD framework, (2) technical platform development, and (3) usability evaluation. Study sites included a large quaternary academic medical center ED and a community hospital ED; all study participants were EM clinicians who practiced at one of these sites.

Ethics Approval

This study was approved by our institutional review board (IRB00185078) after expedited review.

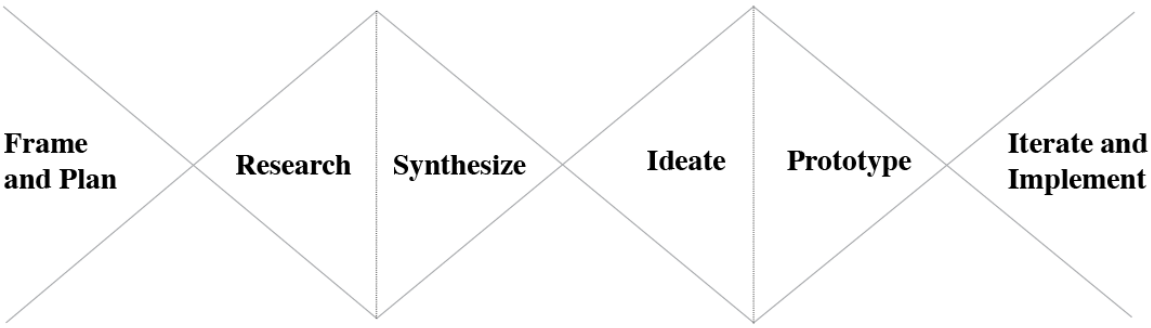
Phase 1: Conceptual Development Under an HCD Framework

In the fall of 2019, we conducted an intensive 16-week HCD studio [21] focused on addressing the delivery of feedback to EM clinicians related to post-ED encounter patient outcomes. MICA design faculty (BS and CM) led the studio in partnership with CDEM researchers. As shown in Figure 1, our HCD studio

consisted of 6 stages: *frame and plan, research, synthesize, ideate, prototype,* and *iterate and implement.*

First, our multidisciplinary research team conducted cocreation sessions to fine-tune the scope and objectives of the project. We then engaged in design research with end users via a mixed methods approach that included observations, semistructured interviews, and surveys of EM clinicians. Observations focused on EM clinician interactions with existing technologies, and semistructured interviews focused on current patient outcome follow-up and practice-based learning behaviors (Table 1). Surveys were used to assess current patient follow-up practices, identify important patient outcomes for post-ED encounter follow-up, and define ideal timeframes for outcome reporting (Multimedia Appendix 1). Thematic analysis of research output was used to synthesize "personas," "design principles," and "opportunity areas" that would guide future HCD studio activities.

Figure 1. Human-centered design methods used to develop Linking Outcomes Of Patients. CDEM: Center for Data Science in Emergency Medicine.



	Methods	Users involved
Frame and Plan August	Core team members reflected on experiences the designers could study	Center for Data Science in Emergency Medicine (CDEM), core data scientists
Research August-October	• Interviews (n=15) • Survey (n=24), Status: Attending • Observation of 2 campuses	Interns, Attendings and Resident Physicians, across Johns Hopkins Hospital and Howard County General
Synthesize October	Analog thematic synthesis; themes comprised of multiple mediums — interviews, survey, and observational data	
Ideate November	Feedback #1: 2-hour long session in which design team prompted participants to sketch out solutions, and act their favorite idea out as a skit.	CDEM directors, core data scientists, physicians, human factors engineer
Prototype November	The design team grouped ideas from shareback to build out dimensionality. Then, created low-fidelity prototypes.	CDEM directors, core data scientists
Iterate and Implement December	Feedback #2: Session in which physicians prioritized features modeled in prototyping into their own dashboard	CDEM directors, core data scientists, physicians

Table 1. Examples of high priority questions from the design research stage interview guide.

Question category	Examples
Rapport building	What does a perfect day at the emergency department look like for you?
Stress	What things currently complicate your decision making in the emergency department? (individual, institutional, environmental)
	Optional follow up: Could you relate that back to feedback?
Feedback	How do you find out what happens to your patients after they leave the emergency department?
	Optional Follow up: Why do you think you don't receive the kind of information/feedback you would like?
Design	What format would prefer to receive feedback on what happens to respiratory infection patients and why? (Prompt: ask about how they prefer seeing info, ie, text, charts, images, video, voice messages)
Diagnostic/decision making	Walk me through how you develop a set of possible diagnoses and how you differentiate between them to come to a final diagnosis.

Next, end users were reincorporated into the HCD studio through an in-person multistakeholder ideation session that also included engineers and data analysts. Previously defined *design principles* and *opportunity areas* were employed as guides, and potential solutions were generated using numerous ideation techniques (eg, brainstorming, brain dumping, sketching, storyboarding). The ideation session output was subsequently used by designers to develop a series of solution prototypes. Prototypes progressed through several levels of fidelity as designers collaborated with clinical and technical research team members to ensure that solutions were compatible with existing health information technology. Finally, designers and end users convened to develop an idealized version of the feedback tool, which incorporated features generated during the ideation and prototyping stages. *User personas* were used as a touch point to build out dimensionality for the ideal feedback tool. Attendees of these sessions also created an *experience map*, which described the envisioned end user's journey with the tool through 4 phases: learning about the tool, using the tool, owning the tool, and implementing the tool into practice.

Phase 2: Technical Platform Development

Patient Outcome Selection

An initial set of patient outcomes was selected for inclusion in the initial version of our tool based on end user preferences (discovered through *design research)* and feasibility of data collection and standardization. Outcomes selected were in-hospital mortality, escalation of level of care (eg, floor to intensive care unit) within 24 hours of ED departure, and return to the ED within 72 hours of discharge. These outcomes are commonly used quality measures [22,23], applicable to the entire ED population, and reliably recorded in the EHR.

Data Capture, Normalization, and Delivery

CDEM data analysts and engineers developed a data processing pipeline to facilitate the population of post-ED encounter patient outcomes within our feedback platform. In brief, raw EHR data were first populated within a Health Insurance Portability and Accountability Act–compliant research computing environment, where a normalization code was developed to identify and label post-ED encounter patient-level events and to attribute patient encounters to individual EM clinicians by using native EHR data fields and timestamps. The normalization code was validated via chart review by an EM clinician and data analyst. Following validation, the normalization code was applied to daily extracts of EHR data within a reporting server, and aggregated views of normalized data were pushed to a presentation server that would power our patient feedback platform.

Digital Feedback Platform Design

Working alongside EM clinicians (JH, CK, and SP) and software engineers (AGS and MT) in a co-design process, our lead design researcher (CM) transitioned the final analog prototype into a fully operational digital platform. Early digital prototypes were built using dummy data and design software that facilitated realistic end-user interaction and rapid iterative improvement (Agile methodology) [24,25]. User needs, including minimum information requirements, optimal outcome definitions, and data labels and data filtering capacities, were further defined within this environment. A final digital mock-up was then used as a template to develop LOOP within the clinician-facing analytic software used by our institution to ensure that our final product adhered to the principles and requirements generated through co-design and would function within our local information technology infrastructure. Throughout this process, our design researcher also worked closely with data analysts and software engineers who led the data processing pipeline development to ensure interoperability of our entire LOOP system.

Phase 3: Usability Evaluation

Usability evaluation was performed by 3 members of the research team: a frontline EM physician (CK), a design researcher (CM), and a health systems engineer with a clinical background (ATS). All participants included were practicing clinicians at an ED study site; members of our study team were excluded from participation in the usability evaluation. Participants were selected using a purposive stratified sample of EM clinicians with representation from multiple end user groups based on trainee status (eg, year of residency), clinical experience (eg, trainee, advanced practice provider, and attending), and gender. After verbal consent was obtained, usability evaluation was performed virtually using an audio-video platform and the sessions were recorded.

Pre-LOOP Survey

Participants first completed a brief anonymous electronic survey. The survey assessed demographics, current method(s) used for

patient outcomes review, baseline knowledge of patient outcomes, and attitudes about their current method(s) of review and outcomes (Multimedia Appendix 2). To assess knowledge as opposed to memory, participants were advised prior to the usability evaluation to bring any materials they use to track their patient outcomes and encouraged to refer to these aids to demonstrate their knowledge about their patient outcomes during the survey, which required participants to estimate the frequency of postencounter patient outcomes over several time periods. We assessed their attitudes about patient outcomes by asking about the confidence in estimates, ease of finding this information, and usefulness of knowing this information. Furthermore, we asked about their willingness to use their current method to find these outcomes, their trust in the data obtained, and whether the information collected is representative of the overall trends for all their patients.

Task Analysis

We then provided participants access to LOOP and performed task analysis while they used the tool. The first set of tasks was determining the number of patients they had seen over the past 30 days who experienced each outcome of interest. Additionally, we asked users to navigate to the chart of a patient who returned within 72 hours and was admitted within our EHR. We also asked them to identify a patient who had died during their hospitalization and email the patient's information to a member of our team (to simulate informing a colleague about a patient outcome). Lastly, we asked the users to locate the list of all patients they had seen in the past 30 days and to identify a patient who was dispositioned to hospital observation. While completing these tasks, participants were asked to use the "think-aloud" method, verbalizing their thought process. Research team observers assessed usability performance metrics such as task completion time and methods of navigation and identified struggle points.

Post-LOOP Survey and Interview

After using LOOP, participants were asked to complete a second survey to assess their knowledge, skills, and attitudes related to using LOOP. This survey asked the same questions as the initial survey and assessed their experience using LOOP. To assess usability of the tool, we combined and adapted the System Usability Scale [16], the Standardized User Experience Percentile Rank Questionnaire [17], and validated instruments composed of statements, and asked users to indicate their level of agreement: strongly disagree, disagree, agree, strongly agree (Multimedia Appendix 2). In the final step of the usability evaluation, we performed semistructured interviews to debrief with the participant about their experience with LOOP related to the perceived benefits, usefulness, and intention to use (Textbox 1). At the end of each usability evaluation, we asked for feedback about their usability evaluation experience, and observers had the opportunity to ask follow-up questions about observations from the task analyses.

Textbox 1. Questions from a usability evaluation semistructured interview.

1. Overall, how would you describe your experience with the Linking Outcomes Of Patients (LOOP)?
2. I know you were asked in the surveys about ease—do you find the interface easy to use? Why?
3. Did you find any aspects difficult? What did you expect to happen?
4. How do you feel about the level of information being shown in the outcomes? Do you find the information easy or difficult to digest? Why?
5. Was there any information you found surprising?
6. Do you find LOOP is more or less effective than your current method of reviewing patients? Why? About how long would you estimate you spend on your current method?
7. Are there other benefits you see to using this tool?
8. When do you see yourself using LOOP? What feature or addition that would bring you back to using LOOP?

Data Analysis

After each usability evaluation session, the research team debriefed, uploaded their data to a secure team folder, and collectively summarized the performance metrics, key findings, and issues to address by comparing notes to reach consensus. Issues raised by users during testing as well their semistructured interviews were classified as either front-end (design) or back-end (data infrastructure) challenges. Additionally, issues were categorized based on benefit to the user (significant benefit/minimal benefit) and effort to address (easy, intermediate, difficult). For statistical analysis of the preusability and postusability evaluation survey results as well as task times, descriptive statistics were calculated and data were visualized using R 4.0.3 (R Core Team).

Results

Phase 1: Conceptual Development Under an HCD Framework

Frame and Plan

Through cocreation sessions, the research team came to consensus on a well-defined focus to create a closed feedback loop with postencounter patient-based outcomes for EM providers. The team also determined the steps for accomplishing the remaining stages of the project, which will be further defined below.

Research and Synthesis

The design researchers completed 18 person-hours of workflow observations across both ED sites and conducted semistructured interviews of 11 EM clinicians. Attending EM physicians,

resident physicians, and advanced practice providers participated in observations and interviews. Surveys were completed by 24 attending EM physicians. Several important themes emerged from observations and interviews, as detailed in Table 2. Many EM clinicians were observed using self-devised work-around solutions to track post-ED encounter patient outcomes, including manual creation of patient lists (electronic or handwritten) to facilitate EHR review in the future and exchange of contact information with patients; others reported similar approaches during interviews. We found that EM clinicians desire information about post-ED encounter patient outcomes and see this type of feedback as important to practice-based learning. They also reported that this information is most often unavailable, and when available, is predominantly negative (ie, associated with adverse patient outcomes). Several clinicians reported that these strategies are only effective when they are time permitted, which is a continuous challenge in the ED environment.

Table 2. Themes from observations and semistructured interviews during the research stage of the human-centered design studio.

Theme	Examples of observations	Examples of interview quotations
Emergency Medicine clinicians value postencounter outcomes–driven feedback.	Physician writes down patient's phone number and reports they do this when they want feedback about that patient's outcome.	*…I wish there was a way they could contact me and say, 'I'm not improving, I'm going to see my primary care provider.'*
	Physician gives their personal telephone number to older patients to enable closed-loop feedback.	*…Post encounter feedback is crucial.*
		(About patient outcome follow-up) *…It's sort of like a vitamin that I have to take every day for my health. It's something that will make me a better doctor in the long run.*
Existing systems for delivery of information about patient outcomes are severely limited.	Physician pulls out the list of patients they keep track of from their pocket. Says they are only able to track a couple of patients in each shift.	*…If patients are not put on a list in (the electronic health record) they disappear.*
		…The feedback (we receive) is not representative of what is actually happening.
Currently available outcomes-driven feedback is predominantly negative.	Physician reports that if something bad happens, clinicians find out from leadership. Physician seems tense when discussing.	*…Feedback is limited to lawsuits and bad outcomes.*
		…I get feedback if someone complains or dies.
Emergency Medicine clinicians use workaround solutions to obtain outcome information for cases perceived as interesting or high risk.	Physician leaves patient notes unsigned so that patients' charts will remain in their electronic health record workflow, forcing additional case review	*…I call patients if I'm concerned about them.*
		…I keep a list of patients on my electronic health record profile when I want to see what happened to them.

These themes were further supported by survey results. The most frequent mode reported for learning about patient outcomes was manual EHR chart review (20 of 24 surveyed EM clinicians), followed by email (7/24), phone (5/24), and face-to-face communications (4/24) between colleagues and patients. A small proportion of EM clinicians (2/24) reported learning about outcomes "haphazardly" and during morbidity and mortality conferences, further reinforcing the idea that most feedback available to EM clinicians is negative. Three-quarters of those surveyed (18/24) preferred to receive both positive and negative feedback, while 6 preferred to receive feedback related to negative outcomes only. Most EM clinicians wanted to know if patients required escalation of care level (eg, transfer from floor to intensive care unit) shortly after admission (20/24), died during their hospitalization (15/24), had a discrepancy between diagnosis assigned at time of admission (ED diagnosis) and diagnosis assigned at hospital discharge (inpatient diagnosis) (14/24) or returned for repeat ED evaluation within a brief time window after ED discharge (16/24). Fewer surveyed clinicians (<30%) wished to be notified when patients filled ED prescriptions, visited an urgent care clinic, followed up with primary care physicians, or had medication treatment regimens changed in the inpatient setting after ED departure.

User personas, guiding design principles, and opportunity areas were also defined using information gathered during design research activities. Six user personas that spanned age groups, learning styles, and affinities for technology were generated and used to drive iterative tool design and development and to perform internal testing of prototypes (Figure 2). Design principles around which all future design and development activities would revolve included (1) recognition and demonstration of value for the clinician's practice, (2) capturing curiosity and encouraging action through knowledge building, (3) prioritization of clear and simple information delivery, and (4) maintenance of flexibility to respond to end user clinicians' needs and preference as they arise throughout the co-design process. Finally, the 4 primary opportunity areas identified for meaningful impact were (1) provision of balanced positive and negative feedback, (2) provision of feedback in a format that allows for improvement of decision-making without overwhelming clinicians, (3) provision of both population-level and patient-level outcome data, and (4) creation of a platform that is customizable at the individual clinician level.

Figure 2. Six user personas generated during human-centered design studios and used to drive tool development and to perform internal prototype testing.

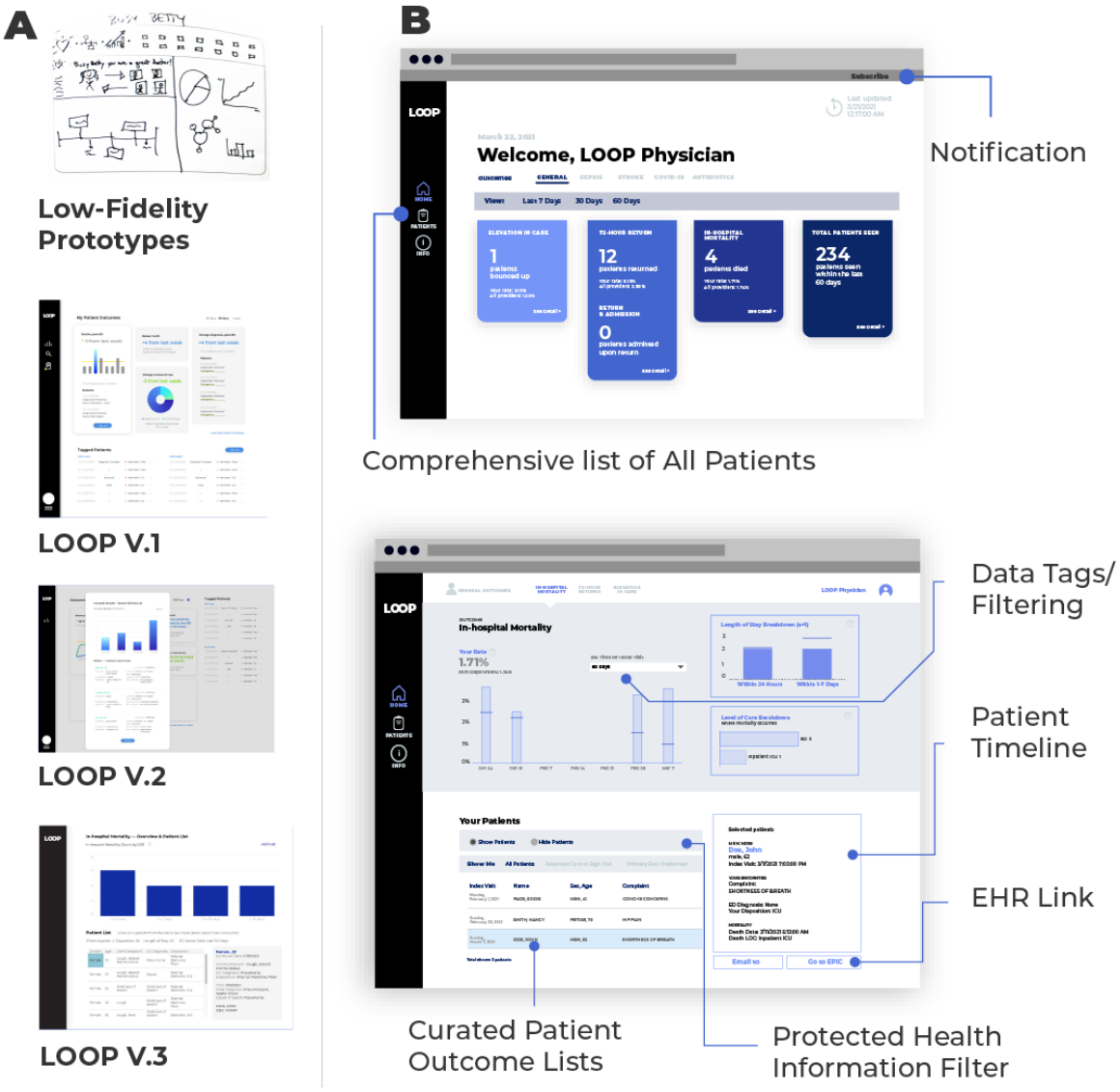
Ideation, Prototyping, and Iteration

During our multistakeholder ideation session, designers collaborated with engineers, data specialists, and end user EM clinicians to translate the *themes*, *opportunity areas*, and *design principles* above into a set of target *design features* that would guide analog prototyping of our tool. Design features that emerged from this session are shown in Table 3. Each participating EM clinician was then assigned a clinician *user persona*, and a 2D sketch representation (first prototype) of a feedback platform was generated for that persona (see Figure 3 for a representative example). Although all design features were represented across the user-generated prototypes, only *comprehensive patient lists*, *data tagging*, and *EHR interoperability* were observed in all prototypes. Over several additional weeks, designers analyzed user prototypes and developed a final analog prototype that included as many design features as possible through iteration.

Table 3. Design requirements established during ideation and prototyping phases of our human-centered design studio.

Design features	Purpose
Comprehensive lists	Allows users to find information of all the patients they have cared for
Electronic health record interoperability	Allows users to transition between platform and patient's clinical chart
Data tagging	Gives users the power to drill down through the use of data labels/tags
Pin a patient	Allows users to prioritize patients of interest for future review
Glow moments	Allows other users to appreciate the work of fellow clinicians
Task timer	Allows a user to customize their experience based on time availability
Home/hospital toggle	Limits visibility of protected health information outside of hospital setting
Patient timeline	Allows users to see a patient's journey after their care
Notification/reminder	Allows users to set an alarm to review outcomes

Figure 3. Evolution of Linking Outcomes Of Patients from human-centered design studios going from (A) physician prototypes to the (B) final version used in the usability evaluation. EHR: electronic health record; LOOP: Linking Outcomes Of Patients.



Phase 2: Technical Platform Development

As shown in Figure 3, analog prototypes were transitioned to a digital feedback platform in a stepwise fashion. The first digital versions of LOOP were developed using design software that was not connected to real-time data flows. Earliest version (Figure 3A) development focused on incorporation of design features defined during phase 1 ideation, while later versions focused on establishing dimensionality that would facilitate attribution of patients to individual clinicians and allow for data sorting and filtering by time and outcome (Figure 3A). Finally, the digital platform was translated into analytic software used by our health care system (Figure 3A) and optimized to accept real-time feeds of normalized EHR data. End user EM clinicians, engineers, and data analysts were included at every stage of digital development. Engineers and data analysts ensured the platform was technically feasible, while end users ensured it was functional and maintained consistency with the themes and design principles generated during our HCD studio. Most design features that were incorporated into early end user protypes (Table 3) were included in the final version of LOOP, including *comprehensive lists*, *EHR interoperability*, *data tagging*, and *home/hospital toggle* (Figure 3B). Other features, including *pin a patient*, *task timer,* and *notification reminder*, were not included as explicit features of the final platform, but tasks associated with these features were possible to perform within the platform using other mechanisms. Others, including *glow moments* and *patient timeline*, were not included in the final platform owing to technical limitations, but they are features that we will seek to incorporate in future versions.

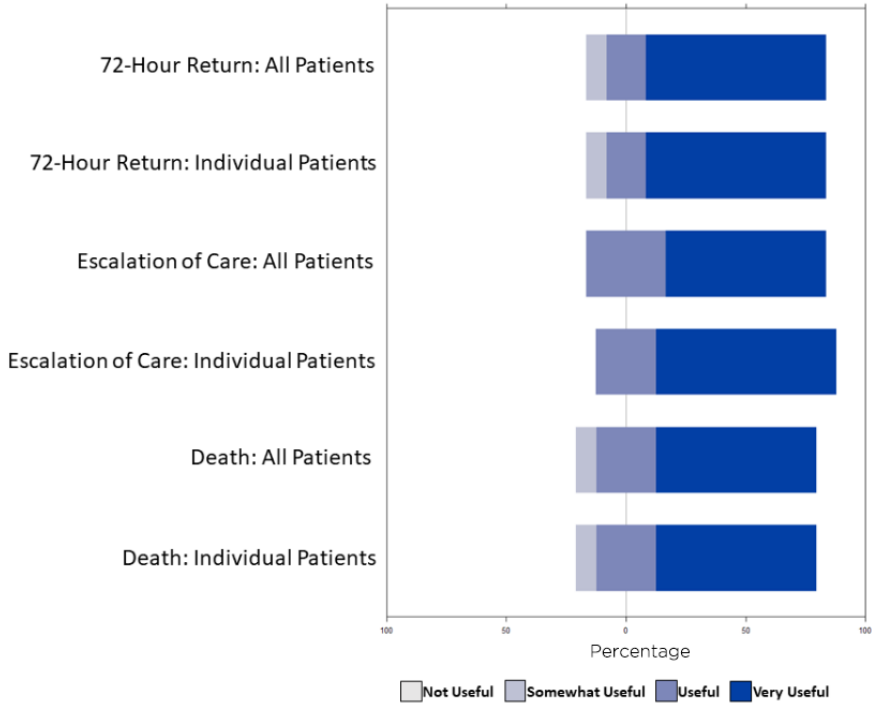
Phase 3: Usability Evaluation

For usability evaluation, our study population included 12 EM providers and 6 (50%) were women. There were 3 (25%) attending physicians, 7 (58%) resident physicians, and 2 (17%) advanced practice providers. The median age was 33.5 (IQR 28-38.3) years. Usability evaluation sessions ranged from 30 to 45 minutes in duration.

Pre-LOOP Survey

When asked about their current method for identifying outcomes for their patients, the median time spent per week to follow-up on patient outcomes was 1.5 (range 0.5-3.5) hours. Of the 12 EM providers, 9 (75%) that described their current method is manually adding each patient to custom-made lists within the EHR; 2 (17%) stated they make handwritten lists of their patients, and 1 (8%) had no method for tracking patient outcomes. Participant attitudes about their current method for determining the outcomes for their patients indicated there was room for improvement. When asked their level of agreement to the statement "I am likely/willing to review my patients using my current method," 11 (92%) users either "strongly disagreed" or "disagreed;" 8 (67%) users disagreed with the statement "I trust the data I am able to find on my patients using my current method." Most users (9/12, 75%) "agreed" or "strongly agreed" that the data gathered using their current method were representative of the overall trends for all their patients. Figure 4 shows participant attitudes about the usefulness of access to outcomes. Most users reported that being able to access all 3 outcomes was "very useful" at the individual patient level and for all their patients.

Figure 4. Participant attitudes about usefulness of access to outcomes at individual patient level and across all their patients.



Task Analysis

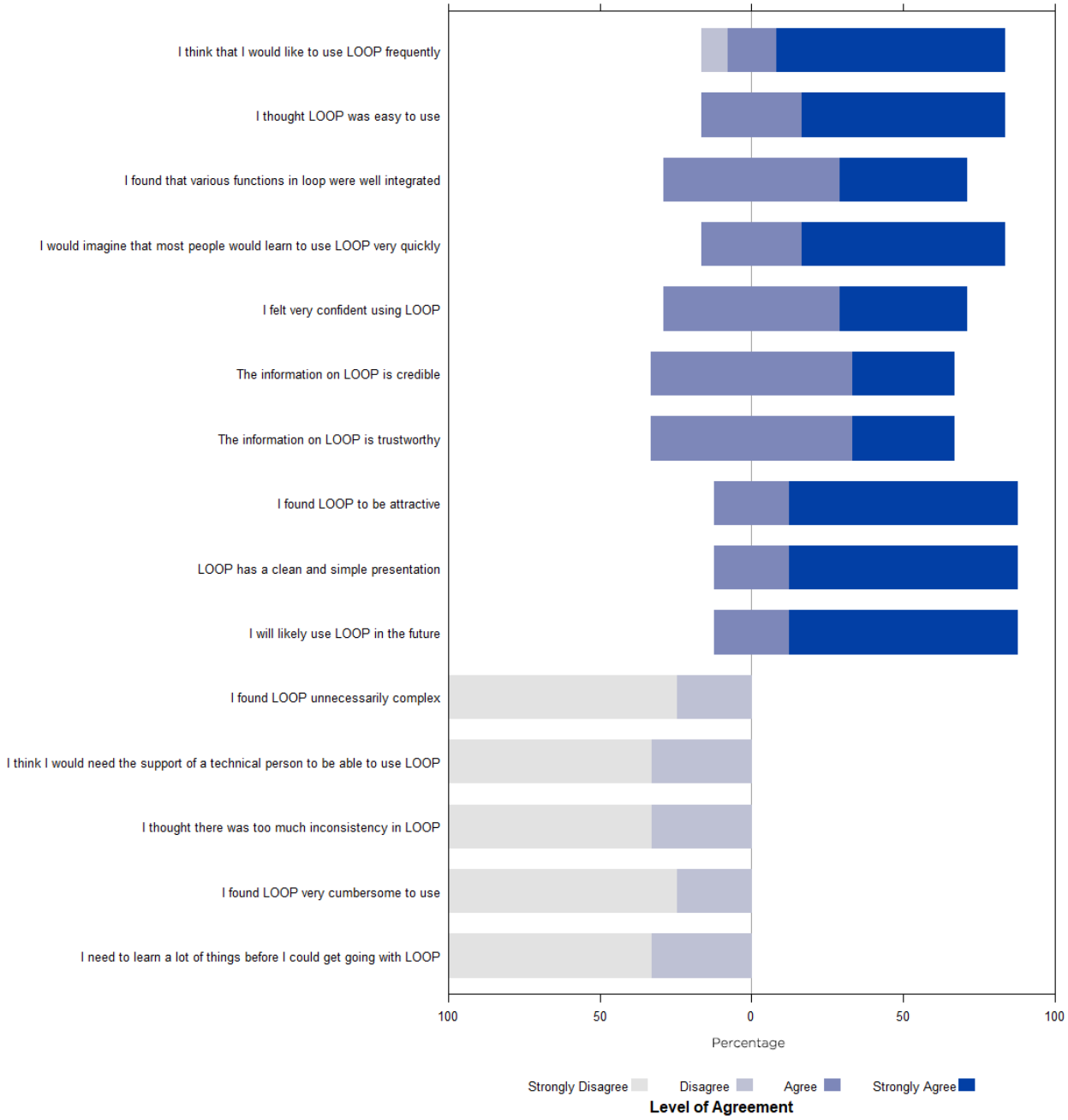
The task analysis of the participants using LOOP to find all 3 outcomes for their patients from the last 30 days was completed in a median of 1.09 (range 0.7-4.3) minutes. The median time to complete all 3 special function tasks (eg, navigating to the electronic record from LOOP, emailing patient information to a colleague, navigating filter features of LOOP) was 6.9 (range 1.7-12.2) minutes. Examples of important observations during this task analysis were (1) the user spent time interpreting a graph instead of noticing the summarizing number somewhere else on the screen and (2) data display errors.

Post-LOOP Survey and Interview

Post-LOOP survey completion time was a median of 1.8 (range 1.1-2.7) minutes. Participants' knowledge of the number of patients that experienced each outcome was different from observed outcomes in LOOP (Figure S1 in Multimedia Appendix 3). Participants underestimated the number of patients who died in hospital but overestimated the number who required an escalation of care outcome. Of note, 1 participant was excluded owing to nonresponse on the initial survey (Figure S2 in Multimedia Appendix 4). The participants' attitudes about their estimates of each patient outcome over the past 30 days improved after using LOOP (Figure S2 in Multimedia Appendix 4). Of note, 1 participant was removed from the analysis for only the question "How easy is it for you to determine this outcome for all your patients?" for all 3 outcomes owing to nonresponse on the post-LOOP survey. For all 3 outcomes, the users shifted from feeling the outcomes were "not easy" to determine at the individual and cumulative patient levels to feeling it was "very easy" after using LOOP. Additionally, they changed from feeling "not confident" and "somewhat confident" about their estimates to overall "very confident" after using LOOP. Participant attitudes about LOOP were overall favorable (Figure 5). On a scale from 0 (unlikely) to 10 (most likely), the users were likely to recommend (score=9.5) LOOP to a colleague. The semistructured interview to debrief with the

users about their LOOP experience helped further inform our understanding of their perceptions about LOOP. The users identified several benefits about LOOP, such as access to data on patients you would not have originally had the time or foresight to follow up on later. One user described LOOP as "much more systematic" than their current methods. Importantly, trainees identified the opportunity to use LOOP as an educational tool that facilitates discussion with their attendings about prior cases with surprising outcomes. A user commented that LOOP is a "fantastic learning tool to make you a better clinician." Regarding the issues identified, the constructive feedback received was helpful and could be addressed. Many concerns centered around harmonizing the visual layout with the functionality on the main page as well as across the platform (eg, connecting summary numbers with the corresponding patient lists). Another issue was improving the defaults and layouts of filtered lists; therefore, the interaction was more intuitive. The ability to discuss with the users while having the tool in front of them to demonstrate the concerns and possible improvements was highly informative.

Figure 5. Participant attitudes about Linking Outcomes Of Patients. LOOP: Linking Outcomes Of Patients.



Discussion

Principal Results

Leveraging previously underutilized EHR data, LOOP was designed and developed to enable systematic delivery of personalized patient outcomes feedback to EM clinicians. The platform allows EM clinicians to (1) quickly review post-ED encounter outcomes at the individual patient level, (2) see outcomes for all patients in their census, and (3) customize views of data based on user preference. Our usability evaluation showed the tool is easy to use and that information presented

within LOOP is viewed as valuable and reliable by end users. The usability evaluation also revealed that information delivered within LOOP is not currently known by EM clinicians. Our team's creation of a tool that is both useful and usable was enabled by a process centered on HCD principles and by a commitment to incorporation of end users at every stage of design and development. Although the version of LOOP reported here includes a selected set of outcomes only (in-hospital mortality, inpatient level of care escalations, and return ED visits), the tool was designed with flexibility to allow for ongoing rapid integration of additional outcomes based on user feedback and clinical need.

Comparison With Prior Work

Absence of patient outcomes feedback in EM is well-documented as is recognition of the importance of this information and a desire to receive it among EM clinicians [6,10,17,26]. Detailed information related to patient outcomes is routinely collected and stored within the EHR data infrastructure; opportunity exists to enable practice-based learning and improvement through standardization and delivery of this information to clinicians [19,20]. However, most efforts to provide patient outcomes feedback in EM have relied on analog and nonsystematic approaches such as manual creation of patient follow-up lists or telephone calls to patients by clinicians [18,27-29]. These approaches are labor-intensive and time-consuming. They also have the potential to promote cognitive bias. When tasked with selecting patients for follow-up themselves, clinicians often focus on cases they already perceived as challenging or at the highest risk for adverse outcome [30,31]. The automated and comprehensive method of data delivery used by LOOP ensures that clinicians receive robust unbiased outcome data, including information that is unexpected and not discoverable using previously described methods. Exposure of these outcomes is important for both early career and seasoned clinicians because unexpected events uncover unconscious deficits and present opportunities to improve competence and increase the quality and safety of care delivered [18,32-34].

Prior efforts to harness EHR data for EM practice improvement have focused on the development of dashboards to guide ED operations management or to enhance real-time display of patient data during the ED encounter [35-37]. To our knowledge, LOOP is the first tool that translates EHR data into post-ED encounter patient outcomes feedback for individual EM clinicians. The systematic delivery of these data by LOOP facilitates deliberate practice in EM, a process whereby expertise can be developed through repeated action and skills improvement, driven by continual feedback and reassessment [7,38,39]. Such practice-based learning is critical to the personal and professional development of clinicians and is mandated by both the ACGME and ABMS [2,3]. The generation and automation of personalized EHR data flows to fill outcome knowledge gaps is a significant step forward for experiential learning in EM. It also represents a step toward more meaningful use of the EHR and development of a learning health care system, which are both the major goals for our nation's health care system [40].

Pragmatic Usability Evaluation

While the potential value of information delivered by LOOP is clear, its real-world utility is dependent on end-user acceptance and long-term adoption. User interface and information display greatly impact whether a tool is adopted by the user [41]. Activities associated with our HCD studio revealed high variation among potential users of LOOP (EM clinicians) with respect to clinical experience, experience and comfort with technology, current practice-based learning behaviors, and desired patient outcomes feedback—all of which were considered during our HCD process. We used a robust and pragmatic approach to assessment that allowed for evaluation of LOOP in a near real-world setting. Our assessment, grounded by the knowledge, skills, and attitudes framework [42], included direct observation and task analysis as users interacted with the tool to find information about real patient encounters, surveys that included standard usability questions and assessed knowledge and attitudes that allowed for comparison with current practices, and semistructured interviews that further explored these topics. We also performed this assessment among a diverse and representative group of end users. Our findings were almost exclusively positive.

Human-Centered Design (HCD)

Our commitment to the use of HCD methods at every stage of this project was critical to its success. The incorporation of user input into the development of information technology platforms is now considered essential in health care [43]. Previously reported clinician user involvement in similar projects is variable, with some groups limiting their involvement to ideation, implementation, or testing phases only [36,37,44]. To our knowledge, the HCD methodology used here is among the most intentional and extensive of those reported to date. This approach was enabled by our multidisciplinary team structure. Longitudinal collaboration between clinician-scientists, engineers, and designers allowed for interpretation of end user contributions through multiple lenses and ensured that the final product of our work was a well-rounded representation of user-generated specifications. We believe this increased end user trust in the final product and will translate to higher rates of adoption in clinical practice. The results of our usability evaluation suggest this is true.

Limitations

This study has several limitations. First, the study was performed within a single health care system, which may limit its generalizability. However, our incorporation of end user EM clinicians throughout design, development, and evaluation activities from varied practice settings (urban academic and suburban community) and levels of training (attending EM physicians, resident physicians, advanced practice providers) minimizes this limitation. In addition, our HCD focus and technical approach to EHR data normalization and presentation are both reproducible and generalizable. Feedback platforms that integrate the needs and wants of clinical staff could be generated by other groups using a similar methodological approach. Second, our usability evaluation was performed in a relatively small sample of EM clinicians. This limitation was minimized through inclusion of a diverse and representative

user group and by inclusion of various qualitative assessment techniques, which included surveys, direct observations, and semistructured interviews. Our sample size was consistent with those previously reported by others and sufficient to reach thematic saturation using these methods [45,46]. Finally, this study did not evaluate long-term adoption rates or the effectiveness of LOOP for clinical practice improvement. These are both important ongoing research objectives of our team. LOOP is currently in use across multiple EDs, and data collection to facilitate study of these questions is underway.

Conclusions

This study demonstrates the potential of HCD in EM and the power of EHR data to augment practice-based learning in episodic care environments. We have outlined a framework for working alongside end-user EM clinicians to develop and test a tool that augments their clinical experience and exposes previously unavailable information to create a closed-loop feedback-driven learning platform. Future objectives include incorporation of additional patient outcomes into LOOP, measurement of long-term adoption rates, and impacts of patient outcomes feedback provided by LOOP on clinical practice.

Acknowledgments

This work was supported by grant R18 HS026640-02 (JH and SL) from the Agency for Healthcare Research and Quality (AHRQ), US Department of Health and Human Services. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of AHRQ. Contributions to this project by members of the digital health company StoCastic (AdA and AD) were supported by grant R18 HS026640-02. We thank all participants in the Connected Emergency Care Patient Safety Learning Lab for their involvement. We also thank Maryland Institute College of Art (MICA)'s codirector of the Center for Social Design, Lee Davis, and MICA's graduate students involved with the design studio: Kadija Hart, Vidisha Agarwalla, Harley French, Sasha Avrutina, Ayangbe Mannen, Eesha Patne, and Eunsoo Kim.

Authors' Contributions

All authors contributed to the editing and final approval of this manuscript. ATS, CM, CEK, BS, SL, and JH contributed to data collection. ATS, CM, CEK, BS, AGS, EK, MT, SL, and JH contributed to data analysis.

Conflicts of Interest

JH serves as StoCastic's Chief Medical Officer and SL serves as StoCastic's Chief Technical Officer; both own equity in StoCastic. This arrangement has been reviewed and approved by the Johns Hopkins University in accordance with its conflict of interest policies. AdA and AD, members of the digital health company StoCastic, also contributed to this project. The other authors do not have any other conflicts of interest.

Multimedia Appendix 1
Survey deployed as part of the human-centered design studio. The survey helped provide more specific data around what outcomes physicians wanted after the encounter for admitted and discharged patients.
[PDF File (Adobe PDF File), 68 KB - humanfactors_v9i1e30130_app1.pdf]

Multimedia Appendix 2
Usability evaluation surveys collecting demographic information, assessing current practice, and capturing pre– and post–Linking Outcomes Of Patients perceptions about their patient outcomes. The post–Linking Outcomes Of Patients survey also specifically addressed their experience with Linking Outcomes Of Patients.
[DOCX File , 32 KB - humanfactors_v9i1e30130_app2.docx]

Multimedia Appendix 3
Participant knowledge about patient outcomes such as (A) in-hospital mortality, (B) escalation of level of care, and (C) return to the emergency department. There were 12 participants, but one is not applicable owing to no response.
[PNG File , 68 KB - humanfactors_v9i1e30130_app3.png]

Multimedia Appendix 4
Participant attitudes about their estimates of pre–Linking Outcomes Of Patients and post–Linking Outcomes Of Patients for patients with each outcome. Three questions were asked for each outcome in the presurvey and postsurvey: 1. How easy is it for you to determine for all your patients? 2. How easy is it for you to determine for an individual patient? 3. How confident are you?
[PNG File , 22 KB - humanfactors_v9i1e30130_app4.png]

References

1. Staker LV. Practice-based learning for improvement: the pursuit of clinical excellence. Tex Med 2000 Oct;96(10):53-60. [Medline: 11070736]
2. Hayden SR, Dufel S, Shih R. Definitions and competencies for practice-based learning and improvement. Acad Emerg Med 2002 Nov;9(11):1242-1248 [FREE Full text] [doi: 10.1111/j.1553-2712.2002.tb01584.x] [Medline: 12414478]
3. Burke AE, Benson B, Englander R, Carraccio C, Hicks PJ. Domain of competence: Practice-based learning and improvement. Acad Pediatr 2014;14(2 Suppl):S38-S54. [doi: 10.1016/j.acap.2013.11.018] [Medline: 24602636]
4. Epstein RM. Mindful practice. JAMA 1999 Sep 01;282(9):833-839. [doi: 10.1001/jama.282.9.833] [Medline: 10478689]
5. Berwick DM. Developing and testing changes in delivery of care. Ann Intern Med 1998 Apr 15;128(8):651-656. [doi: 10.7326/0003-4819-128-8-199804150-00009] [Medline: 9537939]
6. Croskerry P. The feedback sanction. Acad Emerg Med 2000 Nov;7(11):1232-1238 [FREE Full text] [doi: 10.1111/j.1553-2712.2000.tb00468.x] [Medline: 11073471]
7. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004 Oct;79(10 Suppl):S70-S81. [doi: 10.1097/00001888-200410001-00022] [Medline: 15383395]
8. Dine C, Miller J, Fuld A, Bellini L, Iwashyna T. Educating Physicians-in-Training About Resource Utilization and Their Own Outcomes of Care in the Inpatient Setting. J Grad Med Educ 2010 Jun;2(2):175-180 [FREE Full text] [doi: 10.4300/JGME-D-10-00021.1] [Medline: 21975616]
9. Maggard-Gibbons M. The use of report cards and outcome measurements to improve the safety of surgical care: the American College of Surgeons National Surgical Quality Improvement Program. BMJ Qual Saf 2014 Jul;23(7):589-599 [FREE Full text] [doi: 10.1136/bmjqs-2013-002223] [Medline: 24748371]
10. Lavoie CF, Plint AC, Clifford TJ, Gaboury I. "I never hear what happens, even if they die": a survey of emergency physicians about outcome feedback. CJEM 2009 Nov;11(6):523-528. [doi: 10.1017/s1481803500011787] [Medline: 19922711]
11. Juntti-Patinen L, Kuitunen T, Pere P, Neuvonen PJ. Drug-related visits to a district hospital emergency room. Basic Clin Pharmacol Toxicol 2006 Feb;98(2):212-217 [FREE Full text] [doi: 10.1111/j.1742-7843.2006.pto_264.x] [Medline: 16445597]
12. National hospital ambulatory medical care survey: emergency department summary tables. Center for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/nchs/data/nhamcs/web_tables/2016_ed_web_tables.pdf [accessed 2021-01-01]
13. Kovacs G, Croskerry P. Clinical decision making: an emergency medicine perspective. Acad Emerg Med 1999 Sep;6(9):947-952 [FREE Full text] [doi: 10.1111/j.1553-2712.1999.tb01246.x] [Medline: 10490259]
14. Zavala AM, Day GE, Plummer D, Bamford-Wade A. Decision-making under pressure: medical errors in uncertain and dynamic environments. Aust. Health Review 2018;42(4):395. [doi: 10.1071/ah16088]
15. Augustine J. The trip down admission lane. ACEP Now. URL: https://www.acepnow.com/wp-content/uploads/2019/12/ACEP_December_2019.pdf [accessed 2022-03-14]
16. Niska R, Bhuiya F, Xu J. National Hospital Ambulatory Medical Care Survey: 2007 emergency department summary. Natl Health Stat Report 2010 Aug 06(26):1-31 [FREE Full text] [Medline: 20726217]
17. Lavoie CF, Schachter H, Stewart AT, McGowan J. Does outcome feedback make you a better emergency physician? A systematic review and research framework proposal. CJEM 2009 Nov;11(6):545-552. [doi: 10.1017/s1481803500011829] [Medline: 19922715]
18. Chern C, How C, Wang L, Lee C, Graff L. Decreasing clinically significant adverse events using feedback to emergency physicians of telephone follow-up outcomes. Ann Emerg Med 2005 Jan;45(1):15-23. [doi: 10.1016/j.annemergmed.2004.08.012] [Medline: 15635301]
19. Kruse CS, Goswamy R, Raval Y, Marawi S. Challenges and Opportunities of Big Data in Health Care: A Systematic Review. JMIR Med Inform 2016 Nov 21;4(4):e38 [FREE Full text] [doi: 10.2196/medinform.5359] [Medline: 27872036]
20. Janke AT, Overbeek DL, Kocher KE, Levy PD. Exploring the Potential of Predictive Analytics and Big Data in Emergency Care. Ann Emerg Med 2016 Feb;67(2):227-236. [doi: 10.1016/j.annemergmed.2015.06.024] [Medline: 26215667]
21. Melles M, Albayrak A, Goossens R. Innovating health care: key characteristics of human-centered design. Int J Qual Health Care 2021 Jan 12;33(Supplement_1):37-44 [FREE Full text] [doi: 10.1093/intqhc/mzaa127] [Medline: 33068104]
22. Graff L, Stevens C, Spaite D, Foody J. Measuring and improving quality in emergency medicine. Acad Emerg Med 2002 Nov;9(11):1091-1107 [FREE Full text] [doi: 10.1111/j.1553-2712.2002.tb01563.x] [Medline: 12414457]
23. Claeys KC, Lagnf AM, Patel TB, Jacob MG, Davis SL, Rybak MJ. Acute Bacterial Skin and Skin Structure Infections Treated with Intravenous Antibiotics in the Emergency Department or Observational Unit: Experience at the Detroit Medical Center. Infect Dis Ther 2015 Jun;4(2):173-186 [FREE Full text] [doi: 10.1007/s40121-015-0069-7] [Medline: 26055392]
24. Collier K. Introducing agile analytics. In: Agile Analytics: A Value-driven Approach to Business Intelligence and Data Warehousing. Boston, MA: Addison-Wesley Professional; Jul 27, 2011:3-23.
25. What is agile software development? Agile Alliance. 2019. URL: https://www.agilealliance.org/agile101/ [accessed 2020-11-29]
26. Sadosty AT, Stead LG, Boie ET, Goyal DG, Weaver AL, Decker WW. Evaluation of the Educational Utility of Patient Follow-up. Acad Emergency Med 2004 May;11(5):715-719. [doi: 10.1111/j.1553-2712.2004.tb02422.x]
27. Bentley B, DeBehnke D, Ma O. Survey of follow-up systems in emergency medicine residencies: analysis and proposal. Acad Emerg Med 1994;1(2):1116-1120 [FREE Full text] [Medline: 7621161]

28. Gupta R, Siemens I, Campbell S. The use of outcome feedback by emergency medicine physicians: Results of a physician survey. World J Emerg Med 2019;10(1):14-18 [FREE Full text] [doi: 10.5847/wjem.j.1920-8642.2019.01.002] [Medline: 30598713]
29. Wogan JM. Emergency department follow-up of hospitalized patients. Am J Emerg Med 2000 Jul;18(4):510-511. [doi: 10.1053/ajem.2000.7895] [Medline: 10919557]
30. Bowen JL, Ilgen JS, Irby DM, Ten Cate O, O'Brien BC. "You Have to Know the End of the Story": Motivations to Follow Up After Transitions of Clinical Responsibility. Acad Med 2017 Nov;92(11S Association of American Medical Colleges Learn Serve Lead: Proceedings of the 56th Annual Research in Medical Education Sessions):S48-S54. [doi: 10.1097/ACM.0000000000001919] [Medline: 29065023]
31. Bowen JL, O'Brien BC, Ilgen JS, Irby DM, Ten Cate O. Chart stalking, list making, and physicians' efforts to track patients' outcomes after transitioning responsibility. Med Educ 2018 Apr;52(4):404-413. [doi: 10.1111/medu.13509] [Medline: 29383741]
32. McDaniel RR, Jordan ME, Fleeman BF. Surprise, Surprise, Surprise! A complexity science view of the unexpected. Health Care Manage Rev 2003;28(3):266-278. [doi: 10.1097/00004010-200307000-00008] [Medline: 12940348]
33. Vincent C. Understanding and Responding to Adverse Events. N Engl J Med 2003 Mar 13;348(11):1051-1056. [doi: 10.1056/nejmhpr020760]
34. Adams L. Learning a new skill is easier said than done. Gordon Training International. URL: https://www.gordontraining.com/free-workplace-articles/learning-a-new-skill-is-easier-said-than-done/ [accessed 2021-04-29]
35. Mazor I, Heart T, Even A. Simulating the impact of an online digital dashboard in emergency departments on patients length of stay. Journal of Decision Systems 2016 Jun 16;25(sup1):343-353. [doi: 10.1080/12460125.2016.1187422]
36. Franklin A, Gantela S, Shifarraw S, Johnson TR, Robinson DJ, King BR, et al. Dashboard visualizations: Supporting real-time throughput decision-making. J Biomed Inform 2017 Jul;71:211-221 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.024] [Medline: 28579532]
37. Yoo J, Jung KY, Kim T, Lee T, Hwang SY, Yoon H, et al. A Real-Time Autonomous Dashboard for the Emergency Department: 5-Year Case Study. JMIR Mhealth Uhealth 2018 Nov 22;6(11):e10666 [FREE Full text] [doi: 10.2196/10666] [Medline: 30467100]
38. Ericsson K. An expert-performance perspective of research on medical expertise: the study of clinical performance. Med Educ 2007 Dec;41(12):1124-1130. [doi: 10.1111/j.1365-2923.2007.02946.x] [Medline: 18045365]
39. Duvivier RJ, van Dalen J, Muijtjens AM, Moulaert VR, van der Vleuten CP, Scherpbier AJ. The role of deliberate practice in the acquisition of clinical skills. BMC Med Educ 2011 Dec 06;11:101 [FREE Full text] [doi: 10.1186/1472-6920-11-101] [Medline: 22141427]
40. Perlin JB, Baker DB, Brailer DJ, Fridsma DB, Frisse ME, Halamka JD, et al. Information Technology Interoperability and Use for Better Care and Evidence: A Vital Direction for Health and Health Care. NAM Perspectives 2016 Sep 19;6(9):1-16. [doi: 10.31478/201609r]
41. Wright M, Dunbar S, Macpherson B, Moretti E, Del Fiol G, Bolte J, et al. Toward Designing Information Display to Support Critical Care. Appl Clin Inform 2017 Dec 18;07(04):912-929. [doi: 10.4338/aci-2016-03-ra-0033]
42. Kraiger K, Ford J, Salas E. Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology 1993 Nov;78(2):311-328. [doi: 10.1037//0021-9010.78.2.311]
43. Ghazisaeidi M, Safdari R, Torabi M, Mirzaee M, Farzi J, Goodini A. Development of Performance Dashboards in Healthcare Sector: Key Practical Issues. Acta Inform Med 2015 Oct;23(5):317-321 [FREE Full text] [doi: 10.5455/aim.2015.23.317-321] [Medline: 26635442]
44. Swartz JL, Cimino JJ, Fred MR, Green RA, Vawdrey DK. Designing a clinical dashboard to fill information gaps in the emergency department. AMIA Annu Symp Proc 2014;2014:1098-1104 [FREE Full text] [Medline: 25954420]
45. Ryan GW, Bernard HR. Techniques to Identify Themes. Field Methods 2016 Jul 24;15(1):85-109. [doi: 10.1177/1525822x02239569]
46. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough? Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]

Abbreviations

ABMS: American Board of Medical Specialties
ACGME: Accreditation Council for Graduate Medical Education
CDEM: Center for Data Science in Emergency Medicine
ED: emergency department
EHR: electronic health record
EM: emergency medicine
HCD: human-centered design
LOOP: Linking Outcomes Of Patients
MICA: Maryland Institute College of Art

Edited by A Kushniruk; submitted 03.05.21; peer-reviewed by T Johnson, F Ghezelbash; comments to author 28.06.21; revised version received 11.07.21; accepted 11.11.21; published 23.03.22.

Please cite as:
Strauss AT, Morgan C, El Khuri C, Slogeris B, Smith AG, Klein E, Toerper M, DeAngelo A, Debraine A, Peterson S, Gurses AP, Levin S, Hinson J
A Patient Outcomes–Driven Feedback Platform for Emergency Medicine Clinicians: Human-Centered Design and Usability Evaluation of Linking Outcomes Of Patients (LOOP)
JMIR Hum Factors 2022;9(1):e30130
URL: https://humanfactors.jmir.org/2022/1/e30130
doi:10.2196/30130
PMID:35319469

Original Paper

Combining Farmers' Preferences With Evidence-Based Strategies to Prevent and Lower Farmers' Distress: Co-design and Acceptability Testing of ifarmwell

Kate M Gunn[1,2], BPsych (Hons), MPsych (Clin), PhD; Gemma Skaczkowski[1], BHSc (Hons), PhD; James Dollman[3], BSc, MSc, DipEd, PhD; Andrew D Vincent[2], BSc (Hons), PhD; Camille E Short[4], BPsych (Hons), PhD; Susan Brumby[5,6], RN, RM, DipFarm, GDipWSt, MMG, MHlthMgmt, PhD; Alison Barrett[1], MRes, BHlthSC, BIntSt; Nathan Harrison[1], BPsych (Hons); Deborah Turnbull[2,7], BA (Hons), MPsych (Clin), PhD

[1]Department of Rural Health, Allied Health and Human Performance, University of South Australia, Adelaide, Australia

[2]Freemason's Centre for Male Health and Wellbeing, The University of Adelaide, Adelaide, Australia

[3]Alliance for Research in Exercise, Nutrition and Activity, Allied Health and Human Performance, University of South Australia, Adelaide, Australia

[4]Melbourne Centre for Behaviour Change, Melbourne School of Psychological Sciences and Melbourne School of Health Sciences, University of Melbourne, Melbourne, Australia

[5]National Centre for Farmer Health, Western District Health Service, Hamilton, Australia

[6]School of Medicine, Deakin University, Melbourne, Australia

[7]School of Psychology, The University of Adelaide, Adelaide, Australia

Corresponding Author:
Kate M Gunn, BPsych (Hons), MPsych (Clin), PhD
Department of Rural Health
Allied Health and Human Performance
University of South Australia
City East Campus BJ2-03, North Terrace
Adelaide, 5001
Australia
Phone: 61 417852537
Email: kate.gunn@unisa.edu.au

Abstract

Background: Farming is physically and psychologically hazardous. Farmers face many barriers to help seeking from traditional physical and mental health services; however, improved internet access now provides promising avenues for offering support.

Objective: This study aims to co-design with farmers the content and functionality of a website that helps them adopt transferable coping strategies and test its acceptability in the broader farming population.

Methods: Research evidence and expert opinions were synthesized to inform key design principles. A total of 18 farmers detailed what they would like from this type of website. Intervention logic and relevant evidence-based strategies were mapped. Website content was drafted and reviewed by 2 independent mental health professionals. A total of 9 farmers provided detailed qualitative feedback on the face validity of the draft content. Subsequently, 9 farmers provided feedback on the website prototype. Following amendments and internal prototype testing and optimization, prototype usability (ie, completion rate) was examined with 157 registered website users who were (105/157, 66.9%) female, aged 21-73 years; 95.5% (149/156) residing in inner regional to very remote Australia, and 68.2% (107/157) "sheep, cattle and/or grain farmers." Acceptability was examined with a subset of 114 users who rated at least module 1. Interviews with 108 farmers who did not complete all 5 modules helped determine why, and detailed interviews were conducted with 18 purposively sampled users. Updates were then made according to adaptive trial design methodology.

Results: This systematic co-design process resulted in a web-based resource based on acceptance and commitment therapy and designed to overcome barriers to engagement with traditional mental health and well-being strategies—ifarmwell. It was considered an accessible and confidential source of practical and relevant farmer-focused self-help strategies. These strategies were delivered via 5 interactive modules that include written, drawn, and audio- and video-based psychoeducation and exercises, as well as farming-related jokes, metaphors, examples, and imagery. Module 1 included distress screening and information on how to speak

to general practitioners about mental health–related concerns (including a personalized conversation script). Modules were completed fortnightly. SMS text messages offered personalized support and reminders. Qualitative interviews and star ratings demonstrated high module acceptability (average 4.06/5 rating) and suggested that additional reminders, higher quality audio recordings, and shorter modules would be useful. Approximately 37.1% (52/140) of users who started module 1 completed all modules, with *too busy* or *not got to it yet* being the main reason for non-completion, and previous module acceptability not predicting subsequent module completion.

Conclusions: Sequential integration of research evidence, expert knowledge, and farmers' preferences in the co-design process allowed for the development of a self-help intervention that focused on important intervention targets and was acceptable to this difficult-to-engage group.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12617000506392; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=372526

(JMIR Hum Factors 2022;9(1):e27631) doi:10.2196/27631

KEYWORDS

farm; agriculture; rural; drought; mental health; stress; coping; online intervention; acceptance and commitment therapy

Introduction

Background

Farming is an occupation that involves numerous physical and psychological hazards. In recent years, Australian farmers have faced increased exposure to natural disasters, particularly prolonged droughts, fires and floods [1]. Farmers often both live and work on their farms, with family members across multiple generations being involved, consequently blurring the line between work, home, and family roles, which adds to their stress [2-5]. Financial pressure, loss of control, and uncertainty about the future are also associated with environmental stressors and are thought to significantly increase the risk of farmers experiencing mental health problems [3,6,7]. The inability to control these stressors and the sense of hopelessness and entrapment they can engender are thought to be potential risk factors for rural male suicide [8]. Indeed, studies have found a significantly higher incidence of suicide among rural and remote populations compared with metropolitan populations [9,10] and between agricultural workers compared with other employed rural people [11,12].

At the same time, farmers are known to face numerous barriers to help seeking from traditional physical and mental health services. These barriers are structural, such as the limited availability of medical and psychological professionals [13], and attitudinal [14,15]. For generations, Australian farmers have been characterized as being independent, stoic, and skilled at solving practical problems [2,16]. However, in the context of help seeking for the management of psychological distress, traits such as stoicism, independence, and a strong desire to keep personal matters private, may in fact be maladaptive [17]. Recent Australian research has found that farmers were half as likely to have sought help from a general practitioner (GP) or mental health professional in the previous 6 months compared with other employed rural people [13].

Fortunately, the National Broadband Network has now been rolled out in Australia, increasing rural access to internet sites and services [18]. A recent survey of 2000 businesses within the Australian agricultural sector found that up to 95% now have access to the internet [19], and the use of the internet to access health services is known to be increasing in the rural population [20,21].

The delivery of evidence-based interventions on the web offers opportunities to overcome some traditional barriers to help seeking faced by these populations. There is emerging evidence that computerized cognitive behaviour therapy (CCBT) interventions are acceptable in rural communities [22], and an unpublished example of a CCBT intervention designed to address anxiety, depression, and social functioning in Scottish farmers is *Living Life to the Full* (although it reported limited success) [23]. Given farmers' numerous barriers to help seeking and the strong perception within the industry that outsiders (including health professionals [24]) fail to understand their needs and way of life, the development of such interventions needs to be done carefully. Consumer involvement in intervention design ensures that interventions are relevant, usable, and culturally appropriate for the target audience [25,26], which in turn can improve intervention success [27].

Objective

The purpose of this paper is to describe the co-design of content and functionality of a website that aims to help farmers adopt transferable coping strategies that are likely to help them effectively cope with stress. The second purpose of this research is to test the acceptability and feasibility of this website in a broader Australian farming population. The development of this website involved the sequential integration of research evidence, expert knowledge, and farmers' preferences. Methodological guidance and examples such as the studies by O'Brien et al [28] and Short et al [29] and the work outlined in this paper, provide a transparent account of intervention co-design and development upon which other clinicians and researchers can build.

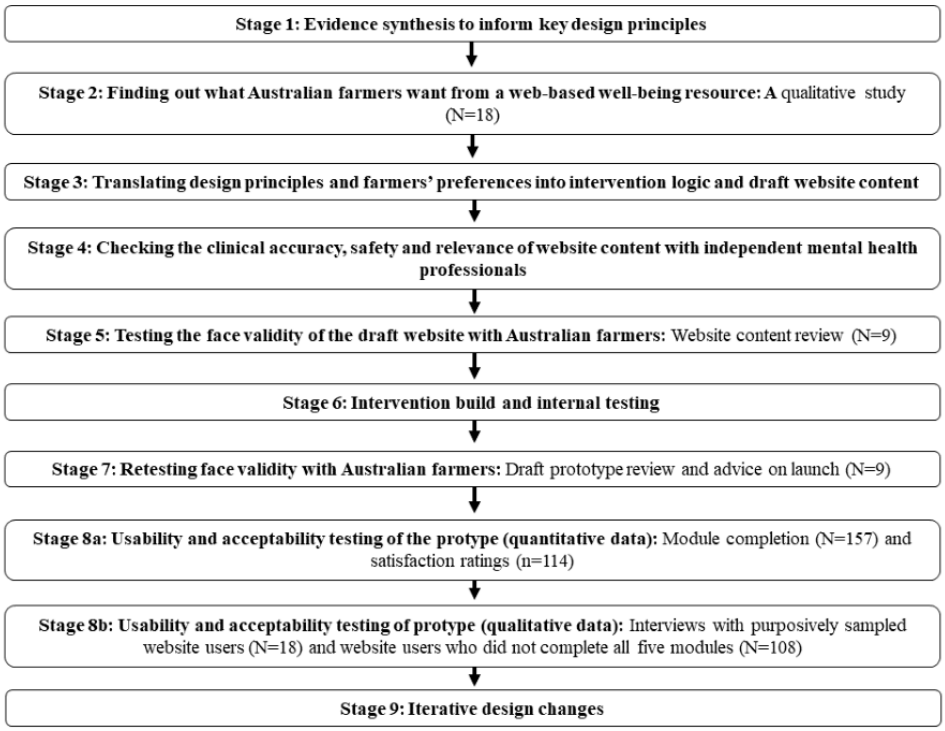
Methods

Overview

Ethics approval for this project was granted by the University of South Australia human research ethics committee (application ID 0000035637). A 9-stage co-design process that included the sequential validation and optimization of evidence and expert

opinion with farmers' wants and preferences was used in a process similar to that described by Easton et al [30]. Each stage resulted in outputs (described in the *Results* section) that were used to inform the next stage of development. Figure 1 summarizes these stages.

Figure 1. Development process for ifarmwell website.



In total, four key methodological approaches informed these stages: (1) synthesis of evidence from prior research to understand the problem and possible solutions (stages 1 and 2); (2) intervention mapping techniques to chart the logic of the intervention (including key acceptance and commitment therapy [ACT] processes or performance objectives, determinants of change, relevant behavior change strategies, and persuasive system design elements; stage 3) [31-33]; (3) a person-based approach via the involvement of farmers as co-designers [34,35] (stages 2, 5, and 7); and (4) iterative updating based on user feedback that allows for ongoing improvements to be made to the website (stages 8 and 9), which is informed by adaptive trial design methodology [36].

All farmers who participated in the research were adults who owned or played an active role in the operation of a farming or pastoral enterprise in Australia (or the spouse of someone who did), were fluent in English, had access to the internet, and had access to a mobile phone with reliable connection or reception at least once per week. The following 9-stage iterative process was conducted over a 3-year period.

Stage 1: Evidence Synthesis to Inform Key Design Principles

Key learnings from published works [6,14,37-40], our own unpublished work, and views from relevant experts across the agricultural, financial, and mental health fields were summarized by the research team. The research team was well-placed to prioritize learnings, given their extensive knowledge of agriculture (KMG, SB, JD, and AB), behavior change interventions (DT, CES, and KMG), web-based interventions (KG, CES, and SB), and rural health (KG, SB, JD, AB, and NH) and mental health (KMG and DT).

Stage 2: Finding Out What Australian Farmers Want From a Web-Based Well-being Resource—A Qualitative Study

Participants

A total of 11 male (11/18, 61%) and 7 female farmers (7/18, 39%), who met the above criteria, participated in the interviews. They had a median age of 45.5 years and were all from grain, sheep, and/or cattle farms across 4 Australian states.

Procedures

As described in detail elsewhere [41], participants were recruited via articles in print, radio and web-based media, advertising via relevant rural organizations, and personal and professional contacts of the research team. Telephone interviews were used to explore the farmers' current internet use practices and preferences for websites designed to promote their mental health and well-being. Thematic analysis was used to analyze the verbatim interview transcripts [42]. Data were arranged under each theme in a Microsoft Excel spreadsheet using a framework approach. The data were checked for any evidence of themes that contradicted the key design principles identified in stage 1.

Stage 3: Translating Design Principles and Farmers' Preferences Into the Intervention Logic and Draft Website Content

The logic of the intervention was systematically developed by KG to ensure that important intervention targets (identified in

stage 1 and explained further in the *Results* section) were addressed and that the effectiveness of the targets could be systematically assessed later. This included mapping the module content to the core ACT processes (acceptance, cognitive defusion, being present, self as context, values, and committed action [43]). It also included ensuring that relevant behavior change techniques [31] (outlined in the *Results* section) were included throughout to help address each of the behavioral determinants (ie, knowledge, skills, emotion, action planning, beliefs about capabilities, beliefs about consequences, motivation and goals, and memory, attention, or decision-making processes) thought to influence whether a user would successfully adopt the core ACT processes. The selection of these behavior change techniques was based on what has been previously shown to effectively address relevant behavioral determinants [31]. Although some overlap with behavior change techniques and persuasive system design elements is acknowledged, persuasive system design elements (as defined by Kelders et al [33] and outlined in the *Results* section) were also built into the intervention logic to help maximize user engagement and limit dropout.

The text, video, and audio content contained within each website module were then drafted by KG by integrating the key design principles from stage 1, farmers' preferences established in stage 2, and the intervention logic identified in stage 3. Her first-hand experience of using ACT in her role as a clinical psychologist, living on a farm in a farming family, developing self-help mental health materials for rural populations, and formal training in intervention mapping, assisted with this process. The general principles of adult learning [44] were also considered.

Stage 4: Checking the Clinical Accuracy and Safety of Website Content With Independent Mental Health Professionals

Participants

A male social worker with a long history of supporting drought-affected farmers and knowledge of and experience using ACT clinically and a female clinical psychologist highly experienced with clinical and forensic mental health populations and in the use of ACT, participated in this stage of testing.

Procedures

Independent feedback on the clinical accuracy, safety, and relevance of website content was provided on all website content using tracked changes in a Microsoft Word processing document. Suggestions were then incorporated where feasible (ie, would not make the modules too long) to enhance clinical impact.

Stage 5: Testing the Face Validity of the Draft Website With Australian Farmers—Website Content Review

Participants

A total of 9 farmers (4/9, 44% men and 5/9, 56% women), who met the criteria outlined above and had participated in stage 2, took part in this stage of the research (herein referred to as *co-designers*). They ranged in age from 34 to 62 years and were from grain, sheep or cattle properties in the states of South Australia (7/9, 78%) and Western Australia (2/9, 22%).

Procedures

A copy of the draft website content was sent to the co-designers via post or email. Participants were also asked to comment specifically on several logo and design options (colors, fonts, background images, and layouts) provided as PDF files. Interviews were then conducted over the phone (or, in one case, in person) to gather feedback, with a focus on language, relevance, and face validity.

Analysis

Where possible, key recommendations for improvement were compiled, and edits were made to the working draft document following the completion of each interview.

Stage 6: Intervention Build and Internal Testing

The purpose of this stage was to produce a working intervention prototype. The research team supplied the website content and design documents developed in earlier steps to a web developer and then worked in close collaboration with them to ensure that lessons from previous stages were integrated into the website and technical glitches were addressed. The prototype was made public in February 2018.

Stage 7: Retesting the Face Validity of the Draft Website With Australian Farmers—Website Prototype Review and Advice on Launch

Participants

A total of 4 farmers (co-designers; 1/4, 25% male and 3/4, 75% female) provided detailed feedback on the website prototype. They were aged 24, 40, 61, and 62 years and were from grain, sheep, and/or cattle properties in South Australia (2/4, 50%), Western Australia (1/4, 25%), and New South Wales (1/4, 25%). A further 5 farmer co-designers (all men) provided feedback specifically on the website launch. They were aged 34, 44, 47, 53, and 55 years and were from sheep or cattle properties (1/5, 20%) or grain, sheep and/or cattle properties (4/5, 80%) in South Australia.

Procedures

Co-designers were sent a link to the website prototype along with broad instructions to work through the website and provide email or phone comments on any aspects they thought required changing.

Analysis

Key recommendations from participant comments were compiled and implemented where possible.

Stage 8a: Usability and Acceptability Testing of the Prototype (Quantitative Data)

The trial was registered with the Australian New Zealand Clinical Trials Registry (ACTRN12617000506392) on April 3, 2017.

Participants

Usability testing of the prototype was conducted by 157 farmers who registered during the study period and met the criteria outlined above. Acceptability testing was undertaken with a subset of 114 users who provided a rating out of 5 for at least module 1. Their demographic characteristics are shown in Table 1.

Table 1. Demographics for all eligible registered users and those users who provided acceptability ratings for at least one module (stage 8a).

Characteristics	All registered users (N=157)	Users who provided acceptability ratings (N=114)
Age (years)		
Values, mean (SD)	45.55 (12.17)	45.46 (12.65)
Values, median (range)	46 (21-73)	46 (21-73)
Gender, n (%)		
Female	105 (66.9)	79 (69.3)
Male	52 (33.1)	35 (30.7)
Remoteness of residence, n (%)[a]		
Major cities of Australia	7 (4.5)	6 (5.3)
Inner regional Australia	66 (42.3)	46 (40.7)
Outer regional Australia	59 (37.8)	43 (38.1)
Remote Australia	16 (10.3)	13 (11.5)
Very remote Australia	8 (5.1)	5 (4.4)
Farm type, n (%)		
Dairy	19 (12.1)	15 (13.2)
Grain, sheep and/or cattle	63 (40.1)	43 (37.7)
Horticulture, market garden, or fruit	14 (8.9)	9 (7.9)
Poultry	3 (1.9)	2 (1.8)
Sheep and/or cattle	44 (28)	37 (32.5)
Viticulture	1 (0.6)	1 (0.9)
Other	13 (8.3)	7 (6.1)
Education level (highest qualification), n (%)		
Postgraduate degree	17 (10.8)	13 (11.4)
University degree or diploma	70 (44.6)	49 (43)
Trade certificate	43 (27.4)	34 (29.8)
Finished high school	25 (15.9)	17 (14.9)
Finished primary school	2 (1.3)	1 (0.9)
Hours per week spent using the internet, mean (SD)	16.42 (10.47)[b]	16.07 (10.23)

[a]n=156 and n=113 because of missing data.

[b]n=155 because of missing data.

Procedures

Consent for participation was established when users registered with the website. Data were collected from all users who registered between February and October 2018 inclusive.

Analysis

Analyses were conducted using SPSS Statistics for Windows (version 26; IBM Corp) [45]. Usability and acceptability were captured in several ways.

Star Ratings (Out of 5) by Each User at the Completion of a Module

At the end of each module, users were asked to rate that module on a scale ranging from 1 to 5 *stars*, where 1=unhelpful, 2=neutral, 3=satisfactory, 4=helpful, and 5=very helpful. The star rating out of 5 was used as it allowed for the multifaceted nature of acceptability to be captured [46] and because of the familiarity and briefness of this approach [47]. Acceptability ratings were examined for modules completed between February and October 2018. Ratings of acceptability for each module were estimated through a linear mixed model with maximum likelihood estimation, and the module number was entered as

a fixed effect with 5 levels and a random intercept per participant. Baseline age, gender, education, farm type, remoteness, hours of internet use, psychological distress, and stress were also entered as fixed factors. The average acceptability rating for each user was calculated from the star ratings of all modules that a user completed.

Module Completion Rate

Data on module completion were captured beyond the February to October 2018 time frame (up to February 2020) to capture participants' full record of participation (even if this was post-October 2018).

Association of Module Completion and Acceptability With Participant Demographics, Recent Exposure to Stressors and Distress Levels

During the registration process, demographics (gender, age, education level, and farm type), distress (Kessler Psychological Distress Scale [48]), and a single-item measure of exposure to stressors were completed. For the latter, users were asked to think of the most stressful situation they had encountered during the past month and rate how stressful they found this situation on a scale of 1 to 10 [40]. Residential postcodes were used to calculate remoteness using the Accessibility and Remoteness Index of Australia from the Australian Bureau of Statistics [49]. Owing to small numbers, the categories *finished high school* and *finished primary school* were combined for analysis. Similarly, poultry farming and viticulture were grouped with *other* farm type.

The association between demographics, stress exposure, distress, and module acceptability was examined using the mixed model described above. A series of univariable and multivariable linear regressions examined the relationship between the number of modules completed and demographic or distress and stress variables. Finally, Pearson correlations were used to examine the association between module completion and an individual's average acceptability rating and the rating of the last module they completed.

Stage 8b: Usability and Acceptability Testing of the Prototype (Qualitative Data)

Participants and Procedures

Brief Phone Calls With Users Who Did Not Complete All 5 Modules (to Find Out Why)

A total of 108 website users who had not continued with the next module within 5 weeks of completing the previous module were followed up with 2 phone calls, 1 email, and 1 additional attempt via email or phone approximately 1 month after that. Successful follow-ups were used to determine the reasons for not continuing with the modules so that we could find ways to enhance the website and aid engagement. Verbatim notes were taken during the phone calls along with email responses, which were manually analyzed by AB and KG using conventional content analysis [50] and a Microsoft Excel spreadsheet. Categories were derived from the data and reworked until all the data could be accounted for. Discrepancies between the coders were rare but were worked through until full agreement was reached.

Detailed Phone Interviews With Purposively Sampled Group of Users

A total of 18 farmers (7/18, 39% men and 11/18, 61% women) who had used the website were purposively selected from website users to gain a variety of impressions (based upon state, farm type, average module acceptability score, gender, and age) and invited via email to take part in a telephone interview to share their experiences. Farmers ranged in age from 23 to 71 years and were from dairy (1/18, 6%), horticulture (2/18, 11%), viticulture (1/18, 6%), sheep and/or cattle properties (7/18, 39%), and grain, sheep and/or cattle properties (6/18, 33%) in Victoria (6/18, 33%), New South Wales (4/18, 22%), South Australia (2/18, 11%), Tasmania (2/18, 11%), Western Australia (2/18, 11%), and Queensland (1/18, 6%). Interviews were audio recorded, transcribed verbatim, and analyzed by AB and KG using thematic analysis [42], with data arranged in a Microsoft Excel spreadsheet using a framework approach, and any discrepancies in coding discussed and reworked until full agreement was reached.

Stage 9: Iterative Design Changes

Following the acceptability assessment of the prototype outlined above, the website was adapted to improve user experience. This aligns with the adaptive trial design methodology [36] and the person-based approach to intervention design of Yardley [34] by continuing to incorporate user feedback after live testing of the intervention.

Results

Stage 1: Evidence Synthesis to Inform Key Design Principles

A summary of our evidence synthesis and the key overarching design principles identified from this are shown in Table 2. In brief, farmers face many barriers to accessing traditional face-to-face mental health services, including a lack of service availability, cost, time, and concerns about confidentiality. They also perceive that outsiders (including health professionals) often do not understand the issues they face. The types of challenges that cause farmers the most stress are those that are beyond their control, and these are the things they feel least equipped to cope effectively with. However, *acceptance* has been shown to be an adaptive coping strategy for farmers in this context [40]. Together, these factors suggest that a new web-based mental health and well-being resource could help overcome existing barriers to engagement by being an accessible, confidential source of farmer-focused, practical self-help strategies based on ACT [51] if co-designed with farmers.

Table 2. Design principles resulting from the evidence synthesis.

Evidence synthesis	Resulting design principle
Barriers to accessing face-to-face mental health and well-being services in rural areas include cost, time, stigma, a lack of anonymity in country towns, a general lack of understanding of mental health issues, and the lack of availability of services [5,16,52-58].	Web-based resources may help to address barriers to the access and availability of services.
Barriers to help seeking for mental health issues among farmers include the desire for control, self-reliance, tendency to minimize the problem, and resignation [14,37]. Farmers prefer anonymous self-help books or internet resources [59].	Self-help resources align with farmers' desire for control, self-reliance, and anonymity.
Farmers are often isolated and perceive a lack of understanding of rural issues from *outsiders* [6,38]. Many farmers report difficulty understanding health care professionals [14] and that health care professionals do not understand them and their way of life [6,53]. However, there is a high level of community trust within rural Australia [39], suggesting that a resource designed by farmers and for farmers may be considered credible.	Having a clear farming focus and co-designing alongside farmers is needed to ensure relevance and acceptability.
Managing uncertainty is a key challenge resulting from drought and a stressor that many farmers do not feel equipped to manage [6]. They are generally already good at solving problems, so they are less likely to benefit from assistance with that.	Uncertainty about the future is a key stressor that farmers need help with managing.
Information provision and educational resources alone are not enough to change key behaviors and thought processes [60]. Evidence-based behavior change techniques (eg, modeling, self-monitoring, and goal setting) should be built into web-based interventions to maximize the effect [33,61].	An interactive, engaging resource is needed.
Farmers who adopt acceptance as a coping strategy and do not engage in behavioral disengagement (giving up) are less likely to experience distress when faced with significant stressors during drought [40].	Acceptance is an effective coping strategy for farmers in this context.
ACT[a] is a transdiagnostic, evidence-based psychotherapeutic approach that can foster acceptance and committed action (opposite of giving up) and improve well-being in a nonpathologizing way [62]. ACT may be used to address a range of psychological disorders and promote general well-being in nonclinical samples [62-64], including via web-based interventions [64,65]. It is particularly suited to contexts where the stressor must be accepted or cannot be fixed [66].	ACT may be an appropriate therapeutic model for this context.
Strategies to improve intervention adherence and effectiveness must also be included (eg, tunneling, personalization, and reminders) [33,67-69].	Issues relating to web-based intervention adherence need to be addressed.

[a]ACT: acceptance and commitment therapy.

Stage 2: Finding Out What Australian Farmers Want From Web-Based Well-being Resources—a Qualitative Study

As reported elsewhere [41], farmers said that they would like a web-based resource that is easy to navigate and compatible with multiple devices and internet connections, as well as their sporadic internet use around work schedules. They preferred a casual and friendly tone, minimal use of jargon, and the inclusion of humor, and they requested information on when and how to seek additional professional help. They also said that they wanted a resource that was authentic, that reflected their challenges and way of life, and that they could see the benefits from quickly. There was no evidence of themes that contradicted the key design principles identified in stage 1.

Stage 3: Translating Design Principles and Farmers' Preferences Into the Intervention Logic and Draft Website Content

Overview

The resulting ifarmwell web-based intervention is a free, farmer-focused, password-protected self-help resource that contains 5 modules. Textbox 1 outlines the purpose of each module as explained to users, and Table 3 details the intervention logic and design, including the key content, targeted ACT processes, behavior change techniques, and persuasive system design elements contained within each module. The content is written for a low reading age (Gunning Fog score=5.8, easily understood by individuals aged 13-14 years) using friendly language with appropriate humor and farming-related metaphors and examples and fits with farmers' ethos of independence and determination to help themselves. The intervention is nonpathologizing and focuses on improving *well-being* and *preventing* poor mental health rather than *treating* poor mental health or mental illness. The word *mental health* is avoided where possible on the website based on farmers' advice about how best to engage their peers.

Textbox 1. ifarmwell module aims (as presented to users).

Module 1: Taking stock of your current well-being and some practical strategies to get you started

- Confidentially discover how your current well-being compares with the well-being of other Australians
- Learn about additional support services that may be useful for you in addition to this web-based resource
- Provide some practical strategies tailored to specific challenges you may face

Module 2: Thoughts are like bullies—how to spend less time *in your head*

- Understand the power thoughts have over the way you feel
- Become more aware of the thoughts or *stories* your mind plays to you
- Learn how to look at your thoughts rather than from them
- Practice evaluating whether a particular thought is helpful to tune in to or not
- Start to learn how to let go of unhelpful thoughts and focus on things that make life better

Module 3: Doing what really matters—how to get the most out of life

- Work out what is important to you
- Identify areas of life in which it would be useful to put more energy
- Recognize areas of your life in which it might be useful to put less energy

Module 4: Training your *attention muscle* and focusing on the *here and now*—a more pleasant, less exhausting place to be

- Become more aware of where your attention is and how this affects how you feel and behave
- Practice shifting attention to the here and now

Module 5: Putting it all together and moving forward

- Revisit strategies
- Plan out how to build these new strategies into day-to-day life
- Think about situations where familiar thoughts or stories may be triggered
- Plan how to respond to new challenges

Table 3. ifarmwell intervention logic and design.

ACT[a] processes	Behavior change techniques (targeting key behavioral determinants of adoption of ACT processes)[b]	Persuasive system design elements (to aid engagement)[c]	Content details
Module 1: taking stock of your current well-being and some practical strategies to get you started			
No ACT processes targeted	• Self-monitoring • Persuasive communication • Information regarding outcomes • Personalized messages • Modeling or demonstration • Goal setting or homework	• Reduction • Tunneling • Tailoring • Personalization • Self-monitoring • Praise • Reminders • Suggestion • Similarity • Liking • Social learning • Normative influence	• Feedback from K10[d] (current levels of distress) and COPE[e] (current coping strategies) • Personalized script for discussion with GP[f] (if medium or high level of distress identified) • Video demonstration of farmer speaking to GP about mental health using a script • Psychoeducation tip sheets for 3 user-identified challenges • Basic self-care and helpful coping strategies (default) • Improving the quality of your sleep • Managing conflict with others • Improving the quality of your relationship • How to get your point across • Managing anger • Coping with grief and loss • Alcohol and drug use • Dealing with domestic violence • Adapting to new roles • What to do if you are feeling down or low • Coping after a natural disaster • Succession planning • Feeling trapped in an unhappy relationship • What to expect in upcoming modules (intro to ACT) • Homework planning or goal setting to implement tip sheet strategies
Module 2: thoughts are like bullies—how to spend less time *in your head*			
• Acceptance • Cognitive defusion • Being present • Self as context (being aware of your experiences without being attached to them)	• Personalized messages • Information regarding outcomes • Self-monitoring • Rewards or positive feedback (encouragement or reinforcement) • Problem-solving • Persuasive communication • Prompts, triggers, and cues • Rehearsal of relevant skills • Graded tasks • Goal setting or homework	• Reduction • Tunneling • Tailoring • Personalization • Self-monitoring • Praise • Reminders • Suggestion • Similarity • Liking • Social learning • Rehearsal	• Homework review or problem-solving obstacles • Feedback from Automatic Thoughts Questionnaire (identification of key challenging stories) • Exploration of existing strategies tried to manage challenging stories. Worked? • Pink sheep or elephants exercise; creative hopelessness • Video: piece-of-paper metaphor demonstration • Audio: notice thoughts while breathing (tool 1) • Examining whether particular thoughts are helpful to focus on or not (drag and drop task with feedback; tool 2) • Drafting thoughts in to just do it, plan a time, and let it go pens • “I’m having the thought that...” exercise (tool 3) • Giving stories a name exercise (tool 4) • Identifying thinking errors (tool 5) • Additional strategies to help you think differently about your thoughts (extra metaphors; tool 6) • Homework planning or goal setting to implement strategies

ACT[a] processes	Behavior change techniques (targeting key behavioral determinants of adoption of ACT processes)[b]	Persuasive system design elements (to aid engagement)[c]	Content details
Module 3: Doing what really matters—how to get the most out of life			
• Values • Committed action	As detailed in module 2 above	• Reduction • Tunneling • Tailoring • Personalization • Self-monitoring • Praise • Reminders • Suggestion • Similarity • Liking	• Homework review or problem-solving obstacles • Consideration of current influences on behavior • Valuing questionnaire and tailored feedback (removed, stage 9) • Values clarification (drag and drop task) • Reflection on current and future decision-making and interactions with others and considering values (tool 7) • Planning to live more consistently with top 10 values in next week and next 6 months (acknowledge what already doing, schedule time, and plan to overcome obstacles; tool 8) • Homework planning or goal setting to implement strategies
Module 4: training your *attention muscle* and focusing on the *here and now*—a more pleasant, less exhausting place to be			
• Being present • Acceptance • Cognitive defusion • Self as context • Values • Committed action	As detailed in module 2 above, plus stress management, relaxation, or mindfulness	• Reduction • Tunneling • Tailoring • Personalization • Self-monitoring • Praise • Reminders • Suggestion • Similarity • Liking • Rehearsal	• Homework review or problem-solving obstacles • Identifying existing activities fully present • Audio: here and now exercise (tool 9) • The basic (mindfulness) formula (tool 10) • Audio: 5 slow, deep breaths grounding technique (tool 11) • Audio: notice 3 things grounding technique (tool 12) • Paying attention to 1 thing at a time when doing everyday activities (tool 13) • Audio: letting go of difficult emotions (tool 14) • Homework planning or goal setting to implement strategies
Module 5: putting it all together and moving forward			
• Acceptance • Cognitive defusion • Being present • Self as context • Values • Committed action	As detailed in module 4 above	• Reduction • Tunneling • Tailoring • Personalization • Self-monitoring • Praise • Reminders • Suggestion • Similarity • Liking • Social learning • Rehearsal	• Homework review or problem-solving obstacles • Audio: leaves on a stream metaphor (tool 15) • Video: normalizes difficulty in mastering these strategies and encourages persistence • Summary of strategies (tool 16) • Audio: cows on a truck metaphor (tool 17) • Relapse prevention (warning signs): how to get yourself back on track and who you could turn to for extra help

[a]ACT: acceptance and commitment therapy.

[b]As defined by Michie et al [31].

[c]As defined by Kelders et al [33].

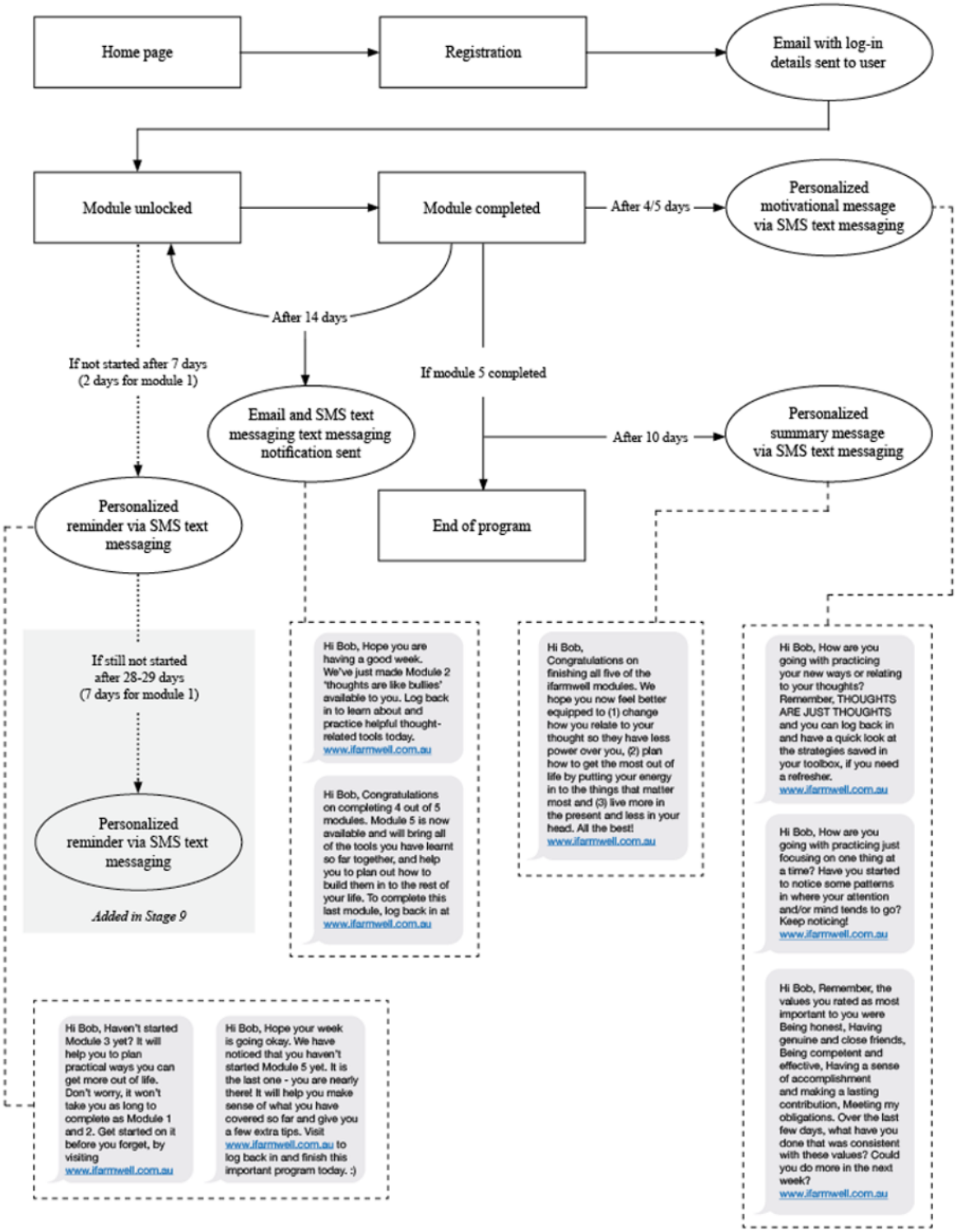
[d]K10: Kessler Psychological Distress Scale.

[e]The COPE inventory [70].

[f]GP: general practitioner.

The intervention was completed over 10 weeks, with each module taking approximately 30 minutes. Users could access the intervention at any time and on any device with an internet connection (eg, laptop, desktop, tablet, and mobile phone). As shown in Figure 2, each module must be completed for the next module to be unlocked. This provided users with time to implement the strategies they learned from the previous module before moving to the next. This design was based on the literature showing that tunneled web-based interventions are less likely to overwhelm users and are better placed to personalize the intervention, leading to greater behavior change [67,68]. Figure 2 also indicates the frequency and type of SMS text messaging reminders sent to users throughout the intervention. Figure 2 shows the final design following changes made after the acceptability testing of the prototype (described in Stage 7).

Figure 2. Wireframe of the ifarmwell website.



Personalization

Tailored content was delivered throughout the intervention based on user responses and demographic variables. This included personalized imagery reflective of participants' farming type, which has also been successfully used in farmer suicide stigma research [69]. In module 1, users were asked to complete the Kessler Psychological Distress Scale measure of distress and were provided with feedback about their current levels of distress, how these compared with others' scores, and inform them if their scores suggested that they should seek professional face-to-face help. More specifically, based on their distress score, users were advised whether they were experiencing what was considered a low (10-19), mild (20-24), moderate (25-29), or severe (30-50) level of distress [71] and subsequently, whether it was recommended that they see their GP to discuss their well-being. Users were given the option of printing off the results of their web-based assessment and a script to guide a conversation about their mental health with their GP. Users were also presented with a short video showing someone else having that conversation with their GP. Finally, any severely distressed users (defined by cut-off) were contacted by a member of the research team by phone or email to encourage them to see a GP and remind them of helpline numbers.

At the end of each module, key tools were summarized, and users could choose to save them to their Toolbox if they found them useful. Modules 2 to 5 contained a homework review component, which asked users about the things they chose to focus on in the previous module and how much they had practiced them since. This section also asked users to reflect on whether anything *got in the way* or made doing this difficult and what they could do in the next week to overcome these difficulties.

At the end of the intervention, the Toolbox provided a summary of the user's existing coping strategies, the stories the user's mind often plays to them, their new preferred tools, and their top values (to guide future decision-making).

Module Content

Module content was transdiagnostic and useful for people experiencing a range of problems or conditions and for people simply wanting to improve their well-being or get more out of life. More specifically, module 1 was designed to take stock of users' current well-being, suggest other sources of help if required, address basic self-care, and provide practical coping strategies that are targeted at users' pressing, unique needs. This was based on a brief suite of questions used to identify the top 3 areas of need for each user. They were then presented with corresponding evidence-based tip sheets (eg, on sleep).

The remaining modules each focused on a particular ACT process. Module 2 addressed the power of thoughts and explained that avoidance or attempts to control difficult thoughts and feelings could be counterproductive. The module asked users to list the emotions and thoughts that they were struggling with, name the stories that they tell themselves, classify them as helpful to focus on, and identify errors in their thinking. Module 3 helped users clarify their values and find ways to live more consistently with them. Module 4 involved several mindfulness-based exercises (not labeled *mindfulness*-based upon farmers' advice), designed to help users identify where their attention was and how this influenced them and practice shifting their attention to the *here and now*. Module 5 summarized the key strategies learned, examined possible triggers and warning signs to be aware of in the future, and reminded users of key sources of support.

Stage 4: Checking the Clinical Accuracy and Safety of the Website Content With Independent Mental Health Professionals

Mental health professionals provided guidance on the appropriateness and safety of the content and suggested minor changes. These included grammatical edits, alterations to simplify the language (eg, *being clear about your values* changed to *knowing what matters to you*), and adding a few more detailed explanations and metaphors to explain key concepts (eg, your mind as an *ideas generator*). Additional reflective questions were also suggested, for example, "what happened to the thought?" following an exercise to help let go of distressing thoughts. It was also recommended that additional text be added to help normalize the fact that one's ability to focus and shift attention may vary from day to day.

Stage 5: Testing the Face Validity of the Draft Website With Australian Farmers—Website Content Review

Overall, participants felt that the module content was acceptable and relevant to farmers. Changes made to the content included repeating icons throughout the modules to guide the user, the inclusion of a summary of content at the beginning of each module, the inclusion of additional cartoons, and the removal of some references to *stress*, which farmers felt their peers would find off-putting (eg, *under pressure* rather than *stressed*). Additional methods for tailoring the content to farmers were also identified. For example, a co-designer suggested likening sorting out thoughts into different categories, to drafting sheep into different pens.

Stage 6: Intervention Build and Internal Testing

A web-based intervention prototype that could be tested by users was created, and wireframes to summarize the website's structure were developed, as detailed in Figure 2. Internal testing by members of the research team resulted in comprehensive lists of hundreds of technical revisions that needed to be made by web developers to improve the user experience.

Stage 7: Retesting the Face Validity of the Draft Website With Australian Farmers—Website Prototype Review and Advice on Launch

This stage resulted in several changes to the look and *feel* of the website, such as a change of font color to improve readability and the inclusion of additional banner photographs featuring machinery and images of younger farmers to ensure broad appeal. Suggestions for improvement also included some website usability issues, such as the ease of saving and returning to a module later. Guidance was also provided on when would be a suitable time of year to advertise and launch the website (ie, not in January when many Australian grain farmers are on

holidays after busy harvests in the lead up to Christmas). The website was made public in February 2018.

Stage 8a: Usability and Acceptability Testing of Prototype (Quantitative Data)

Module Completion Rate

A total of 157 users (described in Table 4) registered on the website between February 2018 and October 2018 and were eligible to participate in the study. Of the 157 users, 17 (10.8%) users registered but did not start module 1. Table 4 shows the total number of people starting and completing each module.

The completion rates for modules 1 to 5 among those that commenced each module were 83.6% (117/140), 89% (81/91), 94% (68/72), 100% (58/58), and 100% (52/52), respectively. Approximately 35% (49/140) of the people who started module 1 did not start module 2 (*dropout*). The dropout rates for modules 2 to 4 were 21% (19/91), 19% (14/72), and 10% (6/58), respectively. Overall, 37.1% (52/140) of the people who started module 1 completed the entire intervention. The median time between starting module 1 and starting module 5 was 16 weeks (8 weeks was intended if users had 2 weeks break before commencing the next module), with a range of 8 to 76 weeks (15/52, 29% took 8-12 weeks; 28/52, 54% took 13-24 weeks; and 9/52, 17% took >24 weeks).

Table 4. Number of users starting and completing each module.

	Started module	Completed module	Completion rate (%)
Module 1	140	117	83.6
Module 2	91	81	89
Module 3	72	68	94.4
Module 4	58	58	100
Module 5	52	52	100

Star Ratings (Out of 5) Provided by Each User at the Completion of a Module

A total of 310 acceptability star ratings were submitted by 114 unique users (those who completed at least module 1 before October 2018). Of 114 users, the average rating across all modules on a 1- to 5-star rating scale was 4.06 (SD 0.99), with 17 (14.9%) people providing an average rating of 1 to 3, 43 (37.7%) people providing an average rating of >3 to 4, and 54 (47.4%) people providing an average rating of >4. The adjusted acceptability ratings for each module from the linear mixed model are shown in Table 5, and the mixed model is shown in Multimedia Appendix 1. There was a significant difference between module ratings; the module 3 acceptability rating was significantly lower than modules 1 (=0.52, 95% CI 0.28-0.77; *P*<.001), 2 (= 0.58, 95% CI 0.32-0.83; *P*<.001), 4 (=0.49, 95% CI 0.20-0.78; *P*=.001), and 5 (=0.75, 95% CI 0.44-1.08; *P*<.001).

Table 5. Adjusted average acceptability ratings (out of 5, where 1=unhelpful and 5=very helpful) for each module.

	Value, mean (SE; 95% CI)
Module 1	4.01 (0.12; 3.77-4.25)
Module 2	4.06 (0.14; 3.79-4.34)
Module 3	3.49 (0.15; 3.19-3.78)
Module 4	3.98 (0.16; 3.67-4.29)
Module 5	4.25 (0.17; 3.91-4.59)

Association of Module Completion and Acceptability With Participant Demographics and Distress Levels

No association was detected between module acceptability and education, farm type, remoteness, age, internet use, or baseline psychological distress (Multimedia Appendix 1). Acceptability ratings were related to stress scores (=0.14, 95% CI 0.06-0.22; *P*=.001); the more stressful the events of the past month, the more satisfied participants were with the modules. Acceptability ratings were related to gender at *P*=.08, indicating a possible trend toward females finding the modules more satisfying than males.

No association was detected between the number of modules completed and gender, education level, farm type, remoteness, hours of internet use per week, or baseline psychological distress or stress exposure (Multimedia Appendix 2). There was an association between module completion and age, with older participants completing more modules (=0.03, 95% CI 0.00-0.06; *P*=.04). Finally, there was no association between module completion and an individual's average acceptability rating (*r*=−0.04; *P*=.52) or their rating of the last module they completed (*r*=0.10; *P*=.28).

Stage 8b: Usability and Acceptability Testing of Prototype (Qualitative Data)

Brief Phone Calls With Users Who Exited the Intervention Before Completion of All 5 Modules (to Find Out Why)

Table 6 summarizes the most frequently identified reasons for not completing a module (N=108). Most often, farmers said they were too busy or simply had not got to it yet (86/108, 79.6%). The next most common reason was that the content was not relevant to them (14/108, 13%) or that they had forgotten about it (8/108, 7.4%).

Table 6. Reasons for not completing all 5 modules (N=108)[a].

Reason	Number of mentions, n (%)
Too busy or not got to it yet	86 (79.6)
Not relevant or helpful for me	14 (13)
I forgot or thought I had done it	8 (7.4)
Technical issues: user end	7 (6.5)
Module took a while or too long	7 (6.5)
Repetitive questions	7 (6.5)
My health	5 (4.6)
Technical issues: ifarmwell	4 (3.7)
Forgot password or reset issues	2 (1.9)

[a]Some participants gave ≥1 reason.

Detailed Phone Interviews With Purposively Sampled Group of Users

In total, 4 broad themes and 25 subthemes were identified and are outlined in Table 7. The themes included the following: using ifarmwell was a positive experience, value for themselves but unsure how best to recommend to others, areas for improvement, and context. The findings generally indicated that users found ifarmwell easy to use and navigate, relevant, credible, and necessary, particularly because of the tough drought conditions that many farmers were experiencing at the time of data collection. Farmers generally liked the structure of the modules and the time provided between modules to practice strategies. They also consistently reported that the language, videos, and cartoons were appropriate, the email or text reminders were helpful, and they valued the opportunity for self-reflection and the anonymity and privacy of the resource. Findings regarding areas for improvement included using even more farmer-focused language, improving the sound quality of the audio files, and including additional reminder SMS text messages to address forgetfulness. Module 3 was also identified as too long, and the values exercises it contained were found to be difficult for people who had never considered this type of value clarification exercise before. The need to double click to answer questions on iPads and iPhones was also something that users said they needed the website to remind them to do.

Table 7. Themes and subthemes from interviews with ifarmwell website users.

Themes	Example quote
Using ifarmwell was a positive experience	• "When I started it off I thought, 'These guys have been reading my mind or watching me,' because it seemed very pertinent, very pertinent. But also, just the fact that there's no shame. I don't have to be ashamed of the fact that I can't help the things I can't help. That's a very empowering and liberating sort of a thing, so I got that from you." [female, 56 years, VIC[a], sheep and/or cattle property]
Easy to use and navigate	• "Very usable, I was really impressed with the usability of it, it was very simple and I am quite computer literate but I can imagine that someone that perhaps wasn't so computer literate, the layout and the sequential nature of it, was pretty good." [female, 61 years, VIC, sheep and/or cattle property]
Relatable and relevant to farmers	• "Yes, if it was just for the ordinary person, which would be, of course, an urban person, it would be very, very different. I'm very grateful it was something focused on farmers because it just - well, it personalised it. It understands what's going on." [male, 64 years, WA[b], grain, sheep and/or cattle farm] • "Because you've structured it for farmers. We're very down-to-earth people, and I think some of these other courses weren't down-to-earth enough. So, your language is being appropriate, your contents are appropriate, illustrations are brilliant." [female, 66 years, NSW[c], horticulture, market garden, and/or fruit growing] • "But like it definitely - yeah read as something that was relevant from a rural perspective and approachable I guess, didn't strike me as someone in an office in Sydney telling us how we should be dealing with the issues of rural mental health or whatever. Like it came across as real." [female, 33 years, QLD[d], sheep and/or cattle property]
Content credible and well-developed	• "I'm just trying to think of the—yeah, I think everything—well, there was nothing in there that I feel was irrelevant or inappropriate at any point." [male, 42 years, NSW, grain, sheep and/or cattle farm] • "Yes, I think it's quite credible. Everything that was written there it was well written, it was easy to understand. I know it said if you need help call Lifeline. There was, that was on there, so yes, it was quite good I thought. Definitely, it looked good. You have obviously spent a lot of work on it. I found it good." [female, 31 years, VIC, grain, sheep and/or cattle farm]
Appropriate language and explanations	• "It was good, it was simple. Not too simple that made you feel like, dumb or anything. They didn't have big words either that you need to look up. So yes, it was quite appropriate I think for the demographic that you're trying to target." [female, 31 years, VIC, grain, sheep and/or cattle farm] • "Just the way you chose the words, you didn't make it more complex than it needed to be and you didn't use technical jargon, it was very simple, everyday language." [female, 61 years, VIC, sheep and/or cattle farm]
Videos were relatable, accessible, good quality, and useful	• Again, I thought they (the videos) were really good because they are relatable and they are real. [female, 55, VIC, Dairy farm]
Appropriate use of photographs and cartoons	• "Yeah, so happy again with those because they really, I think they were chosen well to reflect the environment of the people that you're hoping to reach. You know, kept things within that framework, so yeah, no, absolutely happy with all of those." [male, 25 years, WA, grain, sheep and/or cattle farm]
Modules were presented in a logical sequence	• "I liked the way that it was broken up into different modules so that you were able to look at a section, do the skills and be exposed to some new skills and then have time to consolidate and think about that. For me, that's a really good way to learn new skills, rather than just looking at something on that and then going 'oh that was interesting', it sort of was dribbled out a little bit over a period of time and I found that a really useful format for developing a structure for reflecting on how you deal with life and I think that's a really useful way for a lot of farmers too." [female, 61 years, VIC, sheep and/or cattle farm]
Valued time to implement strategies between modules	• "It was good because it gave you a chance to practice or think about some of the things that you'd discovered, and then—without overloading you, and then you had another follow-up at the next step. I really liked the way that it did that. Like I said, it made it a much more sustainable sort of process." [male, 36 years, VIC, viticulture]

Themes	Example quote
Email, text, or voicemail, exercises, reminders were helpful	• "Yeah, so as much as I hate enlisting in something and they keep bugging me all the time, I thought the texts as well as the things were good, particularly when you've got a fortnight between stuff. Yeah, I thought that was really good." [male, 42 years, SA[e], grain, sheep and/or cattle farm] • "I used to be looking forward to when I got the little message on the phone that said, 'Oi! It's time for you to start doing that extra module.' That's something, I guess, that's important for you lot. The fact that you contact us means that it seems that we've got a relationship or it seems that we matter." [female, 56 years, VIC, sheep and/or cattle farm]
Practical strategies	• "It was good....It was quite practical in the way it was presented, the information was presented...Some of the examples that were presented and things like that were something you can easily identify with. It didn't go into too much detail." [male, 36 years, VIC, viticulture] • "Yeah, I thought it was interesting, I quite like it, I liked the practicality of it actually. I think that was probably the biggest selling point. What I would kind of tell people if I were to recommend it would be there's a lot of practical advice in there, I think that's missing in a lot of stuff. So yeah, no that was definitely the high point of it." [male, 25 years, WA, grain, sheep and/or cattle farm]
Using ifarmwell facilitated self-reflection	• "I think—it took me a long time to identify and realize that I needed to do something with my mental health. It takes a lot to go forward and speak to someone, so being able to go through those modules on your own and identify where you need—you might need some help or even just identifying a few things that you can do for yourself, I think that probably suits farmers or anyone I've ever dealt with at work. I think being able to do something on your own to start with and get a [00:08:38], if this gives you a bit of information, really, to—then if you want to go to someone, you can say, 'Look, this is what I think I need help with.' That's where I really struggled. I didn't know—I didn't really know what to—if I was going to go and talk to someone, I didn't really know what to say. But now I—having gone through those modules, it really highlighted for me." [male, 36 years, VIC, viticulture]
A necessary and timely resource	• "I hope it rolls out because to me it'd be a fair loss if it did not keep going—for sure. So I suppose that means that I better swallow my pride and actually tell someone about it." [male, 42 years, SA, grain, sheep and/or cattle farm] • "I think it's a good program, you'd say, I suppose. It's probably what we need right now too." [female, 31 years, VIC, grain, sheep and/or cattle farm] • "So no, I hope it doesn't disappear because I think there is a definite need there." [female, 33 years, QLD, sheep and/or cattle farm]
Appreciated the opportunity to add tools to Toolbox and refer back to summary sheets over time	• "I quite liked the way you could put stuff in your Toolbox. You could find those things that were potentially going to work for you and put them somewhere so you can refer to them later or coming back to them." [male, 36 years, VIC, viticulture]
Able to understand strategies and apply them to life	• "Just looking at the things I've printed out and stuck on the wall that I thought—be curious. Yes. Always be curious. Always investigate. Put your attention into the here and now. That is—that's important." [female, 71 years, NSW, sheep and/or cattle farm] • "I think it was good. Sure I got some pointers and some tips from that as to how to get over the long and low periods. I mean these are common factors but of course at times when you are down and out you can't think of anything. But these few straight from the program sure help and create that awareness that you can do this or you can do that and give it a go. And it sure help, useful help" [female, 55 years, VIC, sheep and/or cattle farm]
Appreciated privacy and anonymity	• "And like, in this environment, if you want to go to town and go and see a counsellor or a psychologist to say that you happen to live in an area where there is one there, that's probably only going to be an every-now-and-again type visit, it is very difficult to - like I can't even make a doctor's appointment for a script around here with the whole f****** district knowing. Something like that, nobody needs to know. And I know that that actually goes slightly against what we are trying to say is yeah, it is okay to ask for help and it is okay to reach out but sometimes it is actually good to have that first step or offering people resources that doesn't involve anybody knowing about it necessarily." [female, 33 years, QLD, sheep and/or cattle farm]
Willing to recommend to peers	• "Yeah, so I'd definitely be willing to—I reckon it definitely has a space, it fills a need that isn't really getting addressed so far." [male, 25 years, WA, grain, sheep and/or cattle farm] • "That's where I'm doing most of my promoting. I say to the girls—not just girls, to all the people, 'This ifarmwell thing, it was a brilliant idea because this helps. It's particularly tailored for farmers.'" [female, 56 years, VIC, sheep and/or cattle farm]
Value for themselves but unsure about how best to recommend to others	

Themes	Example quote
	• "Yeah I would definitely and I actually thought, while I was going through, there is probably—well, I actually think it would do my partner a lot of good to do it as well, but I haven't quite worked out how to encourage him to do that. But I definitely would given the right type of circumstances" [female, 33 years, QLD, sheep and/or cattle farm]
Areas for improvement	
More farmer-focused language	• "You didn't put enough farmers' language in there." [male, 65 years, SA, sheep and/or cattle farm]
Improve the sound quality of audio files	• "Actually one thing that was a bit of a problem was the, when there was meditations that, the girl that was doing the meditations, her voice was quite low and I couldn't turn it up. So that was a bit of an issue. I could get through with it but it was, that was something that I did notice" [female, 62 years, TAS[f], horticulture, market garden, and/or fruit growing]
Include additional reminders	• "Maybe more reminders. I know for me I obviously signed up and I suppose people that do sign up to do these things do have the intent to do it. Like everything, you sort of get emails from here, there and that's just life these days and that's just the way it is. But I would appreciate obviously another reminder being like 'Come on!'" [female, 23 years, NSW, sheep and/or cattle farm]
Shorten module 3	• "I think that one [module 3] took me the longest time, actually. I did—I think a lot of those things were relevant, and then after a while I dragged and dropped all these things and I began to regret it a little bit, because it took so much time to sort it out and comment on each one. I think that's what happened, so it was a bit lengthy." [male, 64 years, WA, grain, sheep and/or cattle farm]
Remind users to double tap to select answers on iPads and iPhones	• "The only thing- like there was a note about it was that you had to double tap because I did a fair bit of it on my phone and...a couple of times like you would do your multiple choice and I would have to go back because it would say you haven't answered it. I'm like, 'Oh, I did answer it.' But just so obviously hadn't but there was a note in there telling you what you had to do and that was fine but I would say that was more operator error than internet thing." [female, 33 years, QLD, sheep and/or cattle farm]
Context	
Farmers are time poor	• "I am thinking about—from it personally but I am also thinking about it in terms of professionally and how I would perhaps recommend something like that to farmers that I am working with as well and I think that the fact that it's not a very time consuming thing, each module means that you can just do a little bit at a time and you can jump in and out of it, depending on what time requirements you have so the overall structure I thought was terrific from that perspective." [female, 61 years, VIC, sheep and/or cattle farm]
Mental health stigma	• "I guess probably a lot of farmers probably baulk when they hear something about mental health, feelings and emotions and that sort of thing" [male, 36 years, VIC, viticulture] • "I think it's a really good idea because it's—farmers are very proud people. They won't always go and seek help. But this is kind of non-threatening. They don't have to talk to anybody if they don't want to." [female, 57 years, TAS, grain, sheep and/or cattle farm]
Drought	• "We can't do anything about the weather. We can't change it. I haven't got any feed." [female, 71 years, NSW, sheep and/or cattle farm] • "And the other things I liked about it was just that you are farmer-orientated, which is totally different to any of the other help—beyondblue, Black Dog, they're all just for general people but farming situations are particularly unique and your 'ifarmwell' tapped in to that—so the idea that drought or cattle prices that you can't influence and, more importantly, succession." [female, 56 years, VIC, sheep and/or cattle farm]
Women are perceived as most likely to use and recommend	• "I think, the wives, I reckon the wives would be more likely to be interested in it." [female, 62 years, TAS, horticulture, market garden, and/or fruit growing]

[a]VIC: Victoria, Australia.

[b]WA: Western Australia, Australia.

[c]NSW: New South Wales, Australia.

[d]QLD: Queensland, Australia.

[e]SA: South Australia, Australia.

[f]TAS: Tasmania, Australia.

Stage 9: Iterative Design Improvements

In response to the findings detailed above, several changes were made to the website. To improve clarity and brevity, minor wording changes and reductions in the text were made in all modules. Audio recordings were professionally rerecorded to improve quality. Edits to the text were also made to acknowledge that accessing a GP can be difficult for those in rural areas, that module 1, in particular, was very long because of the pre-evaluation questionnaires (but that subsequent modules would involve less reading and more activities), and that questionnaires were standardized and only included for the purposes of website evaluation (not part of the intervention itself; eg, cognitive fusion). To improve usability, additional reminders were included about the inability to *go back* and the need to double tap responses if using an iPad or iPhone. A *Things to remember when using this website* page was added to emphasize these key messages. The *save and continue* button was also made more prominent. To improve relevance, additional images and rotating banners were included on the home page to reflect the broader range of demographics of users accessing the website. To improve adherence, additional SMS text messaging reminders were added 7 days after registration if module 1 was not completed and 28 days after the preceding module was completed if the next module was not started for modules 2 to 5 (Figure 2). Module 3 was shortened by removing 1 value clarification exercise that gave users feedback on values they may not be living consistently with (based on their answers to a questionnaire; Table 3). The revised module 3 retains an exercise asking users to select values that are very important to them, think about whether these values drive their behavior and decision-making, and how they might plan to live more consistently with these values in the future.

Discussion

Principal Findings

This paper outlines the process of integrating evidence from the literature and consumer and expert advice to create a resource that is informed by evidence and perceived as acceptable and relevant by its users. A strength of this intervention development process was the clear, iterative methodology that allowed the integration of different types of knowledge at each step. This involved the synthesis of evidence from prior research and intervention mapping to identify key determinants of behavior change, relevant behavior change and engagement strategies, and the involvement of farmers as co-designers throughout the process to ensure the acceptability of evidence-based strategies. In particular, the farmers' feedback was used to inform the initial design of the website, amend the prototype before launch, inform the timing of the launch, and update the intervention following acceptability and usability testing of the prototype. At all stages, farmers' feedback was prioritized and integrated with research evidence and expert opinions. These approaches enabled us to develop a resource that reflects the unique farming culture, is built on evidence-based approaches to mental health and well-being, demonstrates an understanding of the audience for which it was intended, and as detailed in this paper, was found to be acceptable.

More specifically, the acceptability and usability testing of the prototype that included both quantitative and qualitative components and farmers from a variety of Australian states and farm types, found that once people started a module, most completed that module. Approximately 83.6% (117/140) of users starting module 1 completed module 1, and all people who started modules 4 and 5 completed them. Importantly, acceptability with the previous module was not found to predict whether a user went on to complete the next module, which aligns with the qualitative feedback from people who did not complete every module that their main reason for not progressing was *Too busy/not got to it yet*. Overall, 37.1% (52/140) of people who started module 1 completed the entire 5 module intervention. Comparatively, recent studies have shown a wide variation in the rate of adherence and attrition to web-based interventions for mental health between 2% and 83% [72]. Other studies have reported that approximately 75% do not use mHealth apps more than once after installation [73]. Pleasingly, the present intervention was found to be most acceptable to those who needed it most (ie, those who were most highly distressed when they started module 1) rather than those who were most educated. These high levels of acceptability are significant, given the aforementioned reluctance of farmers to seek help [14], engage with resources targeting their mental health [37], and their general perception that existing services are not designed for them [24]. The intervention also aims to help farmers identify when and how to seek professional help and highlights the role of their local GP. In turn, this may prevent the development of severe mental health problems and facilitate access to treatment at an earlier stage, thereby minimizing the intensity of interventions required and reducing both social and treatment costs. Findings from the qualitative interviews with noncompleters (N=108) to find their reasoning for ceasing participation, also met calls for more research to aid the understanding of engagement in web-based interventions [74] and may be used to inform the inclusion of strategies for improvement in future interventions.

The only comparable farmer-focused well-being website is the aforementioned Scottish CCBT *Living Life to the Full*, which includes personalized support emails in addition to computerized modules [23]. That trial found that of those who logged on (N=35), only 5 (14%) completed the 5 core modules, which is much less than the 37.1% (52/140) reported in this study. Bowyer et al [23] noted that rates of attrition in their study with farmers (73.2%) were much higher than those experienced when they tested a very similar intervention with other population groups (26%-27% attrition) [75,76], reinforcing the notion that the farming population is particularly difficult to engage in health and well-being–focused interventions.

Although acceptability with the ifarmwell modules was generally high, along with the interview comments, they did highlight some areas for improvement. Following the acceptability testing reported in this paper, the website was adapted to address any concerns and improve the user experience. Changes included shortening a module, improving the quality of audio recordings, and incorporating additional SMS text message reminders, which demonstrates the value of adaptive design in building a resource that is responsive to user

experiences. This aligns with the person-based approach by continuing to incorporate user feedback after live testing of the intervention [34], which is a strength of this work as it allows interventions to be responsive to the needs of the audience while remaining publicly available. The need to ensure that modules are as short as possible (or can be easily stopped and recommenced) is important for other farming-focused intervention developers to keep in mind. Our finding that farmers lack the time to engage in web-based interventions aligns with findings that more than half of the Australian farming population work ≥50 hours per week, compared with just 16% of the rest of the working population [77].

Limitations

The sample was limited to those farmers who self-selected to take part in the website evaluation and may not be representative of the wider farming community [78]. Another limitation of this research is that it was not clear to many users that the questionnaires used for evaluation purposes were not part of the intervention itself, which may have contributed to the perception of module length and negatively affected user acceptability. A yellow background was used behind the evaluation components; however, in the future, this delineation should be made even clearer, possibly by having users access the questionnaires via a separate window.

Further Research

We have demonstrated that a co-designed website is usable and acceptable to farmers, and many of the lessons from this research may be applied to the development of future farmer-focused interventions. However, further research is needed to systematically test the effectiveness of this intervention and examine the psychological mechanisms that facilitate changes (or otherwise) in outcomes. In the case of ifarmwell, analyses should specifically examine whether key ACT processes (Table 3) are influenced by the intervention and, if so, how they relate to any changes in distress and well-being outcomes. This would not only inform further refinements to the ifarmwell website but also help progress important gaps in knowledge about psychological mechanisms in the field of ACT [43,51].

Conclusions

This paper describes the first web-based intervention co-designed with farmers to help them adopt coping strategies to better manage their stress by accepting things beyond their control and living according to their values, regardless of the circumstances they face. Importantly, this paper outlines the value of a co-design approach in facilitating the development of interventions that are centered on evidence-based therapeutic approaches, that also appeal to audiences who are typically reluctant to seek help for mental health problems. It also details a comprehensive, successful website development and acceptability testing process, which may inform the development of future web-based interventions for difficult-to-reach populations.

Acknowledgments

The authors gratefully acknowledge the farmers who took part in this research and assisted with the development and testing of the intervention, and Mrs Margaret Mc Gee for her assistance with supporting this project.

This project was funded by the NAB Foundation, University of South Australia, Freemasons Foundation Centre for Men's Health, and the Freemasons Foundation.

Authors' Contributions

KMG was involved in the conceptualization, methodology, funding acquisition, supervision, project administration, data collection, data analysis (qualitative), and writing of the original draft. GS was involved in the data analysis (quantitative) and writing of the original draft. JD was involved in the conceptualization, methodology, funding acquisition, and writing—review and editing. ADV was involved in the conceptualization, methodology, data analysis (quantitative), and writing—review and editing. CES was involved in the conceptualization, methodology, and writing—review and editing. SB was involved in the conceptualization, methodology, and writing—review and editing. AB was involved in the project administration, data collection, data analysis (qualitative), and writing—review and editing. NH was involved in the project administration, methodology, data collection, and writing—review and editing. DT was involved in the conceptualization, methodology, and writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Linear mixed model assessing the relationship between module number, demographic and distress variables on module acceptability rating.
[DOCX File , 15 KB - humanfactors_v9i1e27631_app1.docx]

Multimedia Appendix 2
Univariable and multivariable linear regression models predicting highest number of modules completed.

[DOCX File , 15 KB - humanfactors_v9i1e27631_app2.docx]

References

1. Rainfall deficiencies and water availability. Bureau of Meteorology, Australian Government. 2020. URL: http://www.bom.gov.au/climate/drought/ [accessed 2021-11-30]
2. Fraser CE, Smith KB, Judd F, Humphreys JS, Fragar LJ, Henderson A. Farming and mental health problems and mental illness. Int J Soc Psychiatry 2005 Dec 27;51(4):340-349. [doi: 10.1177/0020764005060844] [Medline: 16400909]
3. Sartore G, Kelly B, Stain H, Albrecht G, Higginbotham N. Control, uncertainty, and expectations for the future: a qualitative study of the impact of drought on a rural Australian community. Rur Rem Health 2008 Aug 8. [doi: 10.22605/rrh950]
4. Gray I, Lawrence G. Predictors of stress among australian farmers. Aus J Soc Issues 2016 Dec 22;31(2):173-189. [doi: 10.1002/j.1839-4655.1996.tb01048.x]
5. Gregoire A. The mental health of farmers. Occup Med (Lond) 2002 Dec;52(8):471-476. [doi: 10.1093/occmed/52.8.471] [Medline: 12488518]
6. Fennell KM, Jarrett CE, Kettler LJ, Dollman J, Turnbull DA. "Watching the bank balance build up then blow away and the rain clouds do the same": a thematic analysis of South Australian farmers' sources of stress during drought. J Rural Stud 2016 Aug;46:102-110. [doi: 10.1016/j.jrurstud.2016.05.005]
7. Sartore G, Kelly B, Stain HJ. Drought and its effect on mental health--how GPs can help. Aust Fam Physician 2007 Dec;36(12):990-993 [FREE Full text] [Medline: 18075619]
8. Alston M. Rural male suicide in Australia. Soc Sci Med 2012 Feb;74(4):515-522. [doi: 10.1016/j.socscimed.2010.04.036] [Medline: 20541304]
9. Farmer J, White C, McCosker A, Perkins D, Dalton H, Powell N, et al. Improving the mental health of rural Australians: a review. Report Phase 1 Prepared by Social Innovation Research Institute, Swinburne University of Technology. 2020. URL: https://apo.org.au/node/303323 [accessed 2021-11-30]
10. Caldwell TM, Jorm AF, Dear KB. Suicide and mental health in rural, remote and metropolitan areas in Australia. Med J Aust 2004 Oct 04;181(S7):10-14. [doi: 10.5694/j.1326-5377.2004.tb06348.x] [Medline: 15462636]
11. Milner A, Spittal MJ, Pirkis J, LaMontagne AD. Suicide by occupation: systematic review and meta-analysis. Br J Psychiatry 2018 Jan 02;203(6):409-416. [doi: 10.1192/bjp.bp.113.128405]
12. Andersen K, Hawgood J, Klieve H, Kõlves K, De Leo D. Suicide in selected occupations in Queensland: evidence from the State suicide register. Aust N Z J Psychiatry 2010 Mar;44(3):243-249. [doi: 10.3109/00048670903487142] [Medline: 20180726]
13. Brew B, Inder K, Allen J, Thomas M, Kelly B. The health and wellbeing of Australian farmers: a longitudinal cohort study. BMC Public Health 2016 Sep 15;16(1):988 [FREE Full text] [doi: 10.1186/s12889-016-3664-y] [Medline: 27634298]
14. Hull MJ, Fennell KM, Vallury K, Jones M, Dollman J. A comparison of barriers to mental health support-seeking among farming and non-farming adults in rural South Australia. Aust J Rural Health 2017 Dec 15;25(6):347-353. [doi: 10.1111/ajr.12352] [Medline: 28618088]
15. Kaukiainen A, Kõlves K. Too tough to ask for help? Stoicism and attitudes to mental health professionals in rural Australia. Rur Rem Health 2020 Mar 2. [doi: 10.22605/rrh5399]
16. Roufeil L, Lipzker A. Psychology services in rural and remote Australia. InPsych: The Bulletin of the Australian Psychological Society Ltd 2007;29(5):8-11 [FREE Full text]
17. Judd F, Jackson H, Komiti A, Murray G, Fraser C, Grieve A, et al. Help-seeking by rural residents for mental health problems: the importance of agrarian values. Aust N Z J Psychiatry 2006 Sep;40(9):769-776. [doi: 10.1080/j.1440-1614.2006.01882.x] [Medline: 16911752]
18. Thomas J, Barraket J, Wilson CK, Cook K, Rennie E, Ewing S, et al. Measuring Australia's Digital Divide: The Australian Digital Inclusion Index 2019. RMIT University, Centre for Social Impact, Telstra. 2019. URL: https://apo.org.au/node/255341 [accessed 2021-11-30]
19. Dufty N, Jackson T. Information and communication technology use in Australian agriculture. Australian Government, Department of Agriculture, Water and the Environment. URL: https://www.awe.gov.au/abares/research-topics/productivity/productivity-publications/ict-use-summary [accessed 2021-11-30]
20. Household use of information technology, Australia, 2014-15. Australian Bureau of Statistics. 2016. URL: https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/8146.0Main+Features12014-15?OpenDocument= [accessed 2021-11-30]
21. Household use of information technology - 2016-17. Australian Bureau of Statistics. 2018. URL: https://www.abs.gov.au/statistics/industry/technology-and-innovation/household-use-information-technology/latest-release [accessed 2021-11-30]
22. Vallury KD, Jones M, Oosterbroek C. Computerized cognitive behavior therapy for anxiety and depression in rural areas: a systematic review. J Med Internet Res 2015 Jun 05;17(6):e139 [FREE Full text] [doi: 10.2196/jmir.4145] [Medline: 26048193]
23. Bowyer HL. An online CBT-based life skills course for the farming community: a feasibility study. D Clin Psy Thesis, University of Glasgow. 2017. URL: https://theses.gla.ac.uk/8402/1/2017BowyerDClinPsy.pdf [accessed 2021-11-30]

24. Brumby SA, Ruldolphi J, Rohlman D, Donham KJ. Translating agricultural health and medicine education across the Pacific: a United States and Australian comparison study. Rur Rem Health 2017 Mar 14;17(1):3931 [FREE Full text] [doi: 10.22605/rrh3931] [Medline: 28292189]
25. Ram M, Grocott PR, Weir HC. Issues and challenges of involving users in medical device development. Health Expect 2008 Mar;11(1):63-71 [FREE Full text] [doi: 10.1111/j.1369-7625.2007.00464.x] [Medline: 18275403]
26. Nilsen E, Myrhaug H, Johansen M, Oliver S, Oxman A. Methods of consumer involvement in developing healthcare policy and research, clinical practice guidelines and patient information material. Cochrane Database Syst Rev 2006 Jul 19;3(3):CD004563 [FREE Full text] [doi: 10.1002/14651858.CD004563.pub2] [Medline: 16856050]
27. Waller A, Franklin V, Pagliari C, Greene S. Participatory design of a text message scheduling system to support young people with diabetes. Health Informatics J 2006 Dec 25;12(4):304-318 [FREE Full text] [doi: 10.1177/1460458206070023] [Medline: 17093001]
28. O'Brien N, Heaven B, Teal G, Evans EH, Cleland C, Moffatt S, et al. Integrating evidence from systematic reviews, qualitative research, and expert knowledge using co-design techniques to develop a web-based intervention for people in the retirement transition. J Med Internet Res 2016 Dec 03;18(8):e210 [FREE Full text] [doi: 10.2196/jmir.5790] [Medline: 27489143]
29. Short CE, James EL, Plotnikoff RC. Theory-and evidence-based development and process evaluation of the Move More for Life program: a tailored-print intervention designed to promote physical activity among post-treatment breast cancer survivors. Int J Behav Nutr Phys Act 2013 Nov 05;10:124 [FREE Full text] [doi: 10.1186/1479-5868-10-124] [Medline: 24192320]
30. Easton K, Potter S, Bec R, Bennion M, Christensen H, Grindell C, et al. A virtual agent to support individuals living with physical and mental comorbidities: co-design and acceptability testing. J Med Internet Res 2019 May 30;21(5):e12996 [FREE Full text] [doi: 10.2196/12996] [Medline: 31148545]
31. Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. Appl Psychol 2008;57(4):660-680. [doi: 10.1111/j.1464-0597.2008.00341.x]
32. Bartholomew-Eldredge L, Markham C, Ruiter R, Fernandez M, Kok G, Parcel G. Planning Health Promotion Programs: An Intervention Mapping Approach. 4th Ed. San Francisco, CA: Jossey Bass; 2016:1-704.
33. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
34. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
35. Yardley L, Ainsworth B, Arden-Close E, Muller I. The person-based approach to enhancing the acceptability and feasibility of interventions. Pilot Feasibility Stud 2015 Oct 26;1(1):37 [FREE Full text] [doi: 10.1186/s40814-015-0033-z] [Medline: 27965815]
36. Chow S, Chang M. Adaptive design methods in clinical trials - a review. Orphanet J Rare Dis 2008;3:11 [FREE Full text] [doi: 10.1186/1750-1172-3-11] [Medline: 18454853]
37. Fennell K, Hull M, Jones M, Dollman J. A comparison of barriers to accessing services for mental and physical health conditions in a sample of rural Australian adults. Rur Rem Health 2018 Feb;18(1):4155 [FREE Full text] [doi: 10.22605/RRH4155] [Medline: 29451985]
38. Gunn K, Turnbull D, McWha JL, Davies M, Olver I. Psychosocial service use: a qualitative exploration from the perspective of rural Australian cancer patients. Support Care Cancer 2013 Sep 2;21(9):2547-2555. [doi: 10.1007/s00520-013-1812-9] [Medline: 23636646]
39. Gunn KM, Berry NM, Meng X, Wilson CJ, Dollman J, Woodman RJ, et al. Differences in the health, mental health and health-promoting behaviours of rural versus urban cancer survivors in Australia. Support Care Cancer 2020 Mar 21;28(2):633-643. [doi: 10.1007/s00520-019-04822-0] [Medline: 31115668]
40. Gunn KM, Turnbull DA, Dollman J, Kettler L, Bamford L, Vincent AD. Why are some drought-affected farmers less distressed than others? The association between stress, psychological distress, acceptance, behavioural disengagement and neuroticism. Aust J Rural Health 2021 Mar 15;29(1):106-116. [doi: 10.1111/ajr.12695] [Medline: 33587319]
41. Gunn KM, Barrett A, Hughes-Barton D, Turnbull D, Short CE, Brumby S, et al. What farmers want from mental health and wellbeing-focused websites and online interventions. J Rur Stud 2021 Aug;86:298-308. [doi: 10.1016/j.jrurstud.2021.06.016]
42. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
43. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006 Jan;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]

44. Lieb S, Goodlad J. Principles of adult learning. HSGD. 2005. URL: http://intranet.hsgd.org/webpage/training%20web%20folders/training%20web/ProgrammedLearning/PaperModules/AdultLearning_Answers.pdf [accessed 2021-12-01]
45. IBM Corp. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corporation; 2019.
46. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
47. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
48. Andrews G, Slade T. Interpreting scores on the Kessler Psychological Distress Scale (K10). Aust N Z J Public Health 2001 Dec;25(6):494-497. [doi: 10.1111/j.1467-842x.2001.tb00310.x] [Medline: 11824981]
49. 1270.0.55.005 - Australian Statistical Geography Standard (ASGS): Volume 5 - Remoteness Structure. Australian Bureau of Statistics. 2016. URL: https://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.005 [accessed 2021-11-30]
50. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
51. Stockton D, Kellett S, Berrios R, Sirois F, Wilkinson N, Miles G. Identifying the underlying mechanisms of change during Acceptance and Commitment Therapy (ACT): a systematic review of contemporary mediation studies. Behav Cogn Psychother 2019 May;47(3):332-362. [doi: 10.1017/S1352465818000553] [Medline: 30284528]
52. Collins JE, Winefield H, Ward L, Turnbull D. Understanding help seeking for mental health in rural South Australia: thematic analytical study. Aust J Prim Health 2009;15(2):159. [doi: 10.1071/py09019]
53. Polain J, Berry H, Hoskin J. Rapid change, climate adversity and the next 'big dry': older farmers' mental health. Aust J Rural Health 2011 Oct;19(5):239-243. [doi: 10.1111/j.1440-1584.2011.01219.x] [Medline: 21933365]
54. Judd F, Cooper A, Fraser C, Davis J. Rural suicide--people or place effects? Aust N Z J Psychiatry 2006 Mar 26;40(3):208-216. [doi: 10.1080/j.1440-1614.2006.01776.x] [Medline: 16476147]
55. Rural and remote health. Australian Institute of Health and Welfare. 2019. URL: https://www.aihw.gov.au/reports/rural-remote-australians/rural-remote-health/contents/summary [accessed 2021-11-30]
56. Complete rethink needed to address rural mental health crisis. National Rural Health Alliance. 2018. URL: https://www.ruralhealth.org.au/news/complete-rethink-needed-address-rural-mental-health-crisis [accessed 2021-12-01]
57. Schirmer J, Yabsley B, Mylek M, Peel D. Wellbeing, resilience and liveability in rural and regional Australia. The 2015 Regional Wellbeing Survey, University of Canberra. 2015. URL: https://core.ac.uk/download/pdf/51344156.pdf [accessed 2021-11-30]
58. Vayro C, Brownlow C, Ireland M, March S. 'Farming is not Just an Occupation [but] a Whole Lifestyle': a qualitative examination of lifestyle and cultural factors affecting mental health help‐seeking in Australian farmers. Sociologia Ruralis 2019 Oct 15;60(1):151-173. [doi: 10.1111/soru.12274]
59. Peck DF, Grant S, McArthur W, Godden D. Psychological impact of foot-and-mouth disease on farmers. J Ment Health 2009 Jul 06;11(5):523-531. [doi: 10.1080/09638230020023877]
60. Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health 2016 Jul;136:109-116 [FREE Full text] [doi: 10.1016/j.puhe.2016.03.030] [Medline: 27184821]
61. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. J Med Internet Res 2011 Mar 14;13(1):e17 [FREE Full text] [doi: 10.2196/jmir.1367] [Medline: 21320854]
62. Brown M, Glendenning A, Hoon AE, John A. Effectiveness of web-delivered acceptance and commitment therapy in relation to mental health and well-being: a systematic review and meta-analysis. J Med Internet Res 2016 Aug 24;18(8):e221 [FREE Full text] [doi: 10.2196/jmir.6200] [Medline: 27558740]
63. Yıldız E. The effects of acceptance and commitment therapy on lifestyle and behavioral changes: a systematic review of randomized controlled trials. Perspect Psychiatr Care 2020 Jul 11;56(3):657-690. [doi: 10.1111/ppc.12482] [Medline: 32043617]
64. Thompson EM, Destree L, Albertella L, Fontenelle LF. Internet-based acceptance and commitment therapy: a transdiagnostic systematic review and meta-analysis for mental health outcomes. Behav Ther 2021 Mar;52(2):492-507. [doi: 10.1016/j.beth.2020.07.002] [Medline: 33622516]
65. Kelson J, Rollin A, Ridout B, Campbell A. Internet-delivered acceptance and commitment therapy for anxiety treatment: systematic review. J Med Internet Res 2019 Jan 29;21(1):e12530 [FREE Full text] [doi: 10.2196/12530] [Medline: 30694201]
66. Hayes SC. Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies - republished article. Behav Ther 2016 Nov;47(6):869-885. [doi: 10.1016/j.beth.2016.11.006] [Medline: 27993338]
67. Pugatch J, Grenen E, Surla S, Schwarz M, Cole-Lewis H. Information architecture of web-based interventions to improve health outcomes: systematic review. J Med Internet Res 2018 Mar 21;20(3):e97 [FREE Full text] [doi: 10.2196/jmir.7867] [Medline: 29563076]

68. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for internet interventions. Ann Behav Med 2009 Aug;38(1):18-27 [FREE Full text] [doi: 10.1007/s12160-009-9133-4] [Medline: 19802647]
69. Kennedy AJ, Brumby SA, Versace VL, Brumby-Rendell T. Online assessment of suicide stigma, literacy and effect in Australia's rural farming community. BMC Public Health 2018 Jul 06;18(1):846 [FREE Full text] [doi: 10.1186/s12889-018-5750-9] [Medline: 29980237]
70. Carver CS, Scheier MF, Weintraub JK. Assessing coping strategies: a theoretically based approach. J Personal Soc Psychol 1989;56(2):267-283. [doi: 10.1037/0022-3514.56.2.267]
71. 4817.0.55.001 - Information Paper: use of the Kessler Psychological Distress Scale in ABS Health Surveys, Australia, 2007-08. Australian Bureau of Statistics. 2007. URL: https://www.abs.gov.au/ausstats/abs@.nsf/mf/4817.0.55.001 [accessed 2021-11-30]
72. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
73. Perez S. Nearly 1 in 4 people abandon mobile apps after only one use. TechCrunch. 2016. URL: https://techcrunch.com/2016/05/31/nearly-1-in-4-people-abandon-mobile-apps-after-only-one-use/ [accessed 2020-11-25]
74. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
75. Espie C, Kyle S, Williams C, Ong J, Douglas N, Hames P, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. Sleep 2012 Jun 01;35(6):769-781 [FREE Full text] [doi: 10.5665/sleep.1872] [Medline: 22654196]
76. Hoyle D, Slater J, Williams C, Schmidt U, Wade TD. Evaluation of a web-based skills intervention for carers of people with anorexia nervosa: a randomized controlled trial. Int J Eat Disord 2013 Sep 25;46(6):634-638. [doi: 10.1002/eat.22144] [Medline: 23712500]
77. Labour force and other characteristics of farmers. Australian Bureau of Statistics. 2012. URL: https://www.abs.gov.au/ausstats/abs@.nsf/lookup/1301.0main+features3032012 [accessed 2021-03-10]
78. Crisp DA, Griffiths KM. Participating in online mental health interventions: who is most likely to sign up and why? Depress Res Treat 2014;2014:790457-790411 [FREE Full text] [doi: 10.1155/2014/790457] [Medline: 24804089]

Abbreviations

ACT: acceptance and commitment therapy
CCBT: computerized cognitive behaviour therapy
GP: general practitioner

Edited by A Kushniruk; submitted 01.02.21; peer-reviewed by K Stephen, P Grace-Farfaglia; comments to author 25.02.21; revised version received 10.03.21; accepted 11.09.21; published 11.01.22.

Please cite as:
Gunn KM, Skaczkowski G, Dollman J, Vincent AD, Short CE, Brumby S, Barrett A, Harrison N, Turnbull D
Combining Farmers' Preferences With Evidence-Based Strategies to Prevent and Lower Farmers' Distress: Co-design and Acceptability Testing of ifarmwell
JMIR Hum Factors 2022;9(1):e27631
URL: https://humanfactors.jmir.org/2022/1/e27631
doi:10.2196/27631
PMID:35014963

Original Paper

Educators' Perspectives on Integrating Technology Into Sexual Health Education: Implementation Study

Martha J Decker[1], MHS, DrPH; Salish Harrison[2], BA; Melisa Price[3], MPH; Abigail Gutmann-Gonzalez[2], MPH; Jennifer Yarger[2], PhD; Rachel Tenney[4], MD

[1]Department of Epidemiology and Biostatistics, Philip R. Lee Institute for Health Policy Studies, University of California, San Francisco, San Francisco, CA, United States

[2]Philip R. Lee Institute for Health Policy Studies, Bixby Center for Global Reproductive Health, University of California, San Francisco, San Francisco, CA, United States

[3]Clinical and Translational Science Center, University of California, Davis, Davis, CA, United States

[4]School of Medicine, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Martha J Decker, MHS, DrPH
Department of Epidemiology and Biostatistics
Philip R. Lee Institute for Health Policy Studies
University of California, San Francisco
550 16th Street, Second Floor
San Francisco, CA, 94158
United States
Phone: 1 (415) 476 3095
Email: Mara.Decker@ucsf.edu

Abstract

Background: In the last decade, the use of technology-based sexual health education has increased. Multiple studies have shown the feasibility of technology-based interventions, while a subset has also shown efficacy in improving youths' sexual health outcomes such as increased condom use and knowledge. However, little is known about health educators' experiences in integrating technology to augment sexual health curricula.

Objective: The purpose of this study was to assess the perceptions and experiences of health educators regarding the incorporation of technology into a sexual health education program designed for underserved youth in Fresno County, California, and to identify facilitators and challenges to incorporating technology into the in-person curriculum.

Methods: This implementation study used data collected as part of a cluster randomized controlled trial to evaluate In the Know (ITK), an in-person sexual health education curriculum that includes technology-based content, such as a resource locator, videos, and games, which can be accessed through a mobile app or website. Data from implementation logs from each cohort (n=51) and annual interviews (n=8) with health educators were analyzed to assess the health educators' experiences using the technology and adaptations made during the implementation.

Results: The health educators reported that technological issues affected implementation to some degree: 87% of the time in the first year, which decreased to 47% in the third year as health educators' familiarity with the app increased and functionality improved. Technology issues were also more common in non–school settings. Successes and challenges in 3 domains emerged: managing technology, usability of the ITK app, and youth engagement. The health educators generally had positive comments about the app and youth engagement with the technology-based content and activities; however, they also noted certain barriers to adolescents' use of the mobile app including limited data storage and battery life on mobile phones.

Conclusions: Health educators require training and support to optimize technology as a resource for engaging with youth and providing sensitive information. Although technology is often presented as a solution to reach underserved populations, educational programs should consider the technological needs and limitations of the participants, educators, and settings.

International Registered Report Identifier (IRRID): RR2-10.2196/18060

(JMIR Hum Factors 2022;9(1):e31381) doi:10.2196/31381

KEYWORDS

adolescent; sex education; technology; mobile app; implementation; California; health educator

Introduction

The use of technology-based sexual health education programs aimed at reducing sexually transmitted infections and unplanned adolescent pregnancy has increased over the last decade. Teaching with technology can be defined as any type of educational process that incorporates digital technology tools such as television, computers, tablets, smartphones, mobile apps, online educational games, or online collaborative learning environments to advance student learning [1].

The use of digital technologies in sexual health education programs has increased for multiple reasons. Some data suggest that youth access to the internet and web-based content has become nearly ubiquitous. A Pew Research report showed that 95% of adolescents aged 13-17 years had access to a smartphone in 2018 with almost 45% reporting being online on a "near-constant" basis and 90% going online multiple times per day [2]. Using technology for entertainment and information seeking may be particularly appealing in adolescence, and technology may also help to reinforce adolescent developmental growth through exploration and social connection [3]. Digital technology may also help alleviate student and teacher embarrassment, which is common when discussing sensitive subjects during sexual health classes [4]. In addition, technological tools may be able to reach and educate marginalized youth who lack access to quality and inclusive sexual health education in their schools [5,6]. However, recent research shows that a digital divide persists even among young people [7]. For example, in 2019, low-income adolescents were less likely to own laptops and smartphones than high-income adolescents (36% vs 54% and 74% vs 89%, respectively) [8].

Prior research has demonstrated that youth have favorable opinions of technology-based sexual and reproductive health interventions [9-12]. Some studies also have shown that interventions that incorporate technology were effective in improving youths' sexual health outcomes, such as condom use, abstinence, sexual health knowledge, and safer sex norms [13-15]. However, a previous review of sexual health education apps found that most lacked comprehensive sexual health content and had limited interactivity, highlighting the unmet potential for this type of platform [16].

Despite this increase in digital sexual health interventions, little is known about health educators' experiences delivering sexual health interventions that incorporate technology-based components. Previous research on technology in general educational programming found that health educators' lack of confidence and perceived value of the technology can be barriers to integration [17,18]. One implementation evaluation of an online sexual health program in the Netherlands reported that while teachers appreciated the interactive content, they often needed to adapt the materials based on classroom dynamics, and some found transitioning between web-based and classroom teaching challenging [9]. Coaches in a sports-based HIV program in South Africa, which included text messages as part of the intervention, identified students' shortage of cellular data as the primary challenge [19]. With the growing interest in online and technological approaches to education, it is critical to learn from the experiences of health educators in incorporating technology to ensure that digital content is a viable resource for engaging with youth and improving sexual health knowledge and behavioral outcomes.

The purpose of this study was to assess the perceptions and experiences of health educators regarding the integration of technology into a sexual health education program called In the Know (ITK) and to identify facilitators and challenges to incorporating technology into the in-person curriculum. These results can help future program developers and health educators anticipate and mitigate common issues with technological integration and promote best practices.

Methods

Intervention Overview

ITK was developed by and for adolescents aged 13-19 with a goal of increasing use of contraceptive and clinical health services [20]. Adolescents representing diverse priority populations engaged in a user-centered design process to help create the program's content and digital components [21,22]. The curriculum is based on a positive youth development approach, which promotes personal strengths and healthy development through supportive opportunities and experiences [23,24].

The program was developed to be inclusive and to address the needs of homeless and unstably housed youth; youth of color; and lesbian, gay, bisexual, transgender, and queer or questioning (LGBTQ+) youth. ITK combines 6 hours of in-person sexual health education with technology-based content to provide the skills, information, and resources necessary to improve the sexual and reproductive health and overall well-being of adolescents.

The intervention is divided into three modules: (1) sexual health and contraceptive use; (2) healthy relationships; and (3) educational and career success. Health educators incorporated different technology-based components in each module, such as videos, online goal setting, career opportunities, and geo-location of local services. Some content was "gamified" using Kahoot, a game-based learning platform, and app-based quizzes and activities to earn points. Health educators concluded each module with a guided activity on the app and then assigned a task for the youth to complete outside of class. Youth could also sign up to receive text message reminders of key content and personal goals. These tools as well as additional resources and quizzes were available through a downloadable app or website, enabling youth to access the information outside of the in-person sessions. Health educators provided tablets with the app previously installed for use during the in-person sessions, though the participants also were encouraged to download the app on their mobile phones. The health educators helped to

troubleshoot any technical issues youth were experiencing with the app.

The health educators received training on the curriculum, classroom management, and the technological components prior to implementation. This included in-person trainings and "teach backs" as well as shorter refresher trainings throughout the implementation period. The health educators reported technical issues about the app to the website developers. In addition, the developers updated features and replaced broken links over the course of the program. The researchers, health educators, and the app developers met biweekly to discuss any implementation challenges and adaptations.

Setting and Participants

ITK was implemented in 51 cohorts (groups) with 559 youth at 36 youth-serving agencies representing a variety of settings where youth receive services or activities in Fresno County, California. The health educators traveled to the sites of the participating agencies for implementation, which included school and after-school settings, employment and training sites, youth development centers, clubs, foster care sites, housing authorities, tribal agencies, and LGBTQ+ programs. The majority of participating youth were Latino (70% [n=381]) with a mean age of 15.5 years (SD=2.07). Almost all of the participants owned or shared a smartphone (89% [n=480]), and 86% (n=469) had access to the internet in their homes.

Over the 3 years of implementation, a total of 6 health educators implemented ITK, with an average of 3 health educators per year. The health educators had a range of educational backgrounds, prior teaching, or training experience, and were comfortable with technology. This varied from 1 educator with over 6 years teaching comprehensive sexual health education to 2 educators with no prior experience in sexual or reproductive health and limited familiarity with technology; 2 other health educators had at least 2 years of experience implementing sexual health education in similar settings. Moreover, 2 health educators were male, and all lived in Fresno County.

Data Collection

This implementation study used data collected as part of the cluster randomized controlled trial [20]. Due to the complexity of the intervention being evaluated, a better understanding of the contextual factors, including the technology and in-person implementation, can help to improve future interventions and interpret the intervention's outcomes [25]. Process data from implementation logs and annual interviews with health educators were collected to assess fidelity to the intervention and to promote ongoing quality improvement.

Implementation Logs

Health educators completed an implementation log after delivering the program to each cohort. A cohort is a distinct group of participants receiving ITK at a specific time, such as a classroom of students. Each log consisted of 6 main sections: physical space, teaching methods, learning environment, youth participation, classroom management, and technology. The health educators were encouraged to comment on any contextual factors or circumstances that facilitated or hindered program delivery for specific activities or for the entire cohort. Each log also included a closed-ended question, "Thinking about what happened across all of the modules of this cohort, how often did technology issues impact implementation?" with the response options being all, most, some, or none of the time. At the end of each cohort, the health educator uploaded the completed log to Box, a secure online file management system. The researchers reviewed the implementation logs for completeness and accuracy after submission and debriefed with the implementing health educator.

Health Educator Interviews

The researchers conducted annual interviews near the end of each school year with the health educators for 3 years. Due to staffing changes over that time, 2-3 health educators were interviewed each year, with 2 of the health educators interviewed twice. Topic areas included implementation experiences, youth reactions, perspectives on the digital technology components, and recommendations. The interviews were conducted in a private office and averaged 53 minutes in length. All interviews were audio recorded and transcribed verbatim. Health educators received a US $20 gift card in appreciation of their time.

Analysis

This study used a modified form of grounded theory in which a set of potential concepts were identified and coded, and additional themes were inductively identified from the data [26]. The qualitative analysis was guided by structural themes based on key areas of research interest, such as technology use, emerging themes from the review of transcripts, and the open-ended responses in the implementation logs [27]. This mixed coding system combined an initial list of codes using the main research questions and additional codes that were added based on further review [28].

One researcher coded all transcripts, while another double-coded a subset and reviewed coding for intercoder consistency. The coded interviews had an interreliability score of 0.80. The research team met regularly to review the coding process, clarify codes, and update the codebook. As needed, the researchers reviewed the quotes that were coded differently and jointly agreed to their coding. The codes were analyzed for patterns, with relevant themes extracted. The findings were also compared by year and by health educator to assess if experiences varied over time or by person. The qualitative coding was conducted using Dedoose, version 8.0.35 (SocioCultural Research Consultants, LLC) [29].

The responses to the closed-ended question regarding the frequency of technology-related interruptions were extracted and summarized using Stata 16 (Stata Corp). We used the Fisher exact test to compare the responses by whether the cohort received the program in the first year of implementation and in a school setting. One-sided *P* values are reported.

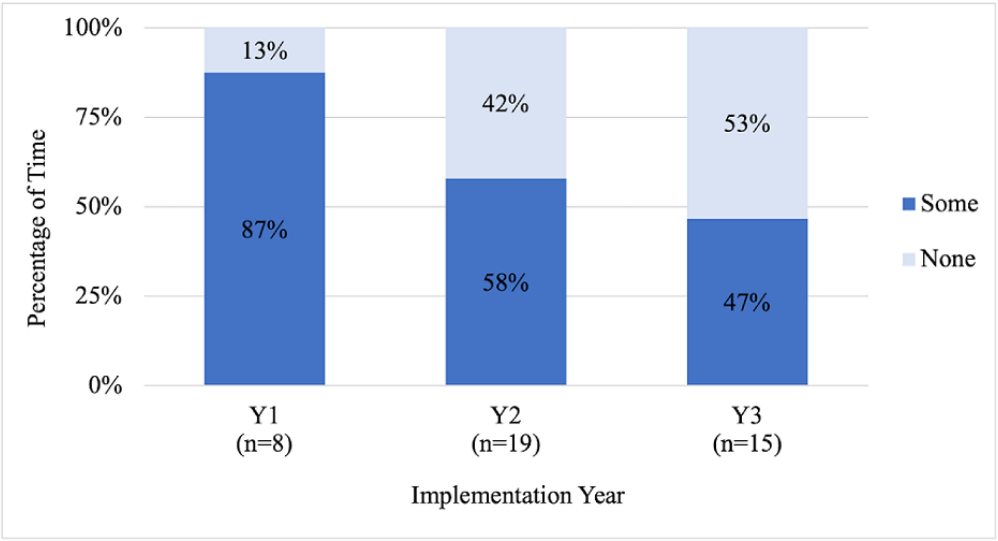
Results

Technology Issues During Implementation

Implementation logs were completed for all 51 in-person sessions of ITK conducted between October 2017 and February

2020. During this time, 8 interviews were completed with the health educators. During the first year of implementation, the health educators reported that technology issues affected implementation to some degree of the time in 7 out of 8 cohorts (87%) with that amount decreasing over the next 2 years, to 11 cohorts out of 19 (58%) and 7 cohorts out of 15 (47%), respectively (Figure 1; note that there were missing responses from 9 implementation logs in year 1 since one question on technology issues was added later). When calculated with the Fisher exact test, the difference between the first year and subsequent years was only marginally significant (*P*=.08) due to the small sample size. The cohorts implemented in non–school settings such as in group homes or community-based organizations were much more likely to have technology issues than those in school settings; 14 out of the 19 (74%) non–school setting cohorts experienced technology issues compared to 11 out of 23 (48%) of cohorts implemented in a school setting (*P*=.05).

Figure 1. Percentage of time when implementation was affected by technology, by year (n=42).



Successes and challenges emerged in 3 key domains: managing technology, usability of the ITK app, and youth engagement. Managing technology included issues related to meeting the technological requirements and administrative needs for implementation during the in-person ITK sessions, such as device compatibility, internet access, and availability of necessary technology hardware. The topics related to the ITK app's usability were those specific to the content and functionality of the app, the integration of the app into the in-person ITK curriculum, and the participants' use of the app. Youth engagement referred to how integrating technology into the curriculum affected the participants' focus and engagement during in-person implementation. Note that many of the issues overlapped; for example, challenges with internet connectivity limited access to the app's content, which then affected youth engagement. Table 1 summarizes the successes and challenges within these 3 domains.

Table 1. Key successes and challenges of integrating technology into in-person sexual health education, by domain.

Domain	Successes	Challenges

Managing technology	• Implementation sites with audiovisual projection devices present (eg, TV, projector, speakers)
	• Bringing mobile Wi-Fi hotspot
	• Providing tablets for classroom use
		• Significant preparation time required
		  • Packing tablets, Wi-Fi packs
		  • Ensuring all devices were charged and functional
		• Implementation site lacked necessary hardware
		• Technology issues caused delays or omission of instruction
		  • Tablets freezing and needing to restart
		  • Internet connectivity issues
		• Inability to connect to the Internet led to alternative instructional methods
		  • Use of hard copies instead of digital content

ITK (In the Know) app usability	• Positive response to online content and resources • App used as reference for local services	• Certain smartphone operating systems not compatible with the app • Specific functionalities within the app repeatedly crashed • Broken links within the app • Barriers to downloading the app on smartphone included lack of data, shared phone, limited battery life • Reluctance to download the app • Youth forgot email address or password needed to access the app • Youth did not use the app after the in-person sessions
Youth engagement	• Use of tablets increased youth engagement • Certain digital content resonated well with youth	• Youth playing on electronic devices led to distractions

Managing Technology

Implementation of ITK involved managing multiple technological devices and administrative requirements such as connecting program tablets and participant smartphones to the internet and projecting digital content on a screen. The health educators noted that implementation was easier in sites that had the necessary audiovisual projecting devices, such as a monitor and projector and internet access. The health educators consistently reported the challenges associated with using technology during implementation, though this decreased in frequency each year. An inability to access high-speed wireless internet was the most commonly described technology issue reported. The health educators adapted to this issue by bringing their own mobile Wi-Fi hotspots to the sites. Other common technology issues included tablets freezing or crashing during use, the lack of audiovisual projecting devices, broken web links to external online content, and youth forgetting log-in information. One health educator described common experiences with the technology:

> *As much as you rely on it and as great as it is, sometimes the links aren't working, the buttons aren't working, the screen goes blank, and you're pressing the button and nothing's working. Then you have to restart it.* [Interview, Year 1]

These technology-related issues caused delays and required health educators to adapt how they delivered the program, both ad hoc and while preparing for future implementation sessions. One health educator described an example of an ad hoc adaptation as such:

> *I had to use downloaded version of materials due to internet connections. Students were really excited for Kahoot [online learning platform] but, unfortunately, the game was not showing the possible answer to the students, and they could not participate the way it is usually played. I ended up reading the questions out loud and had the youth raise hands when the answer sounded correct.* [Log, Year 2]

The health educators provided tablets for participants to access the app if they did not download it on their personal phones. While this increased access to the materials, managing the tablets required significant planning and preparation time, as health educators needed to ensure that all electronic devices were charged and functioning properly. One health educator gave the following explanation:

> *We always get our materials ready… very important is coming to make sure all the tablets are charged… so we don't have any delays the next day.* [Interview, Year 3]

ITK App Usability

The health educators generally had positive comments about the content of the ITK app, particularly the interactive map linking youth to resources within their community. The health educators used these online resources with the participants to identify community clinics, counseling services, and help lines for youth and families experiencing violence. One health educator described the benefits of having resources consolidated on the app:

> *I really do love the resources of [the app]. I always let the youth know, like, "Hey, in the app that we*

> *talked about, you know, if you have any other questions, it'd be really great for you to go on the app and you can find basically anything... There's numbers, there's addresses...” Because a lot of, some of, them do have questions that sometimes I don't know how to answer right off the bat. So I say, “Hey look at the app,” so that's really great.* [Interview, Year 3]

While youth could access the ITK website on a tablet during class, ITK originally was designed as an app for youth to download on their phones for later access. However, youth often faced challenges in using the mobile app including limited storage or data, limited battery life, lack of a personal cell phone number, difficulty remembering the required password, and sharing the phone with other family members. One health educator noted the following experience:

> *You have another group of kids that have phones, but there's always a reason why they don't want to download the app. They don't want to, they don't have space on their phone, their phone is like some crazy off-brand they can't find it. They don't have battery, phone is totally cracked, something about service, something about problems downloading the app. I don't know, it's different every time.* [Interview, Year 2]

Youth also expressed reluctance to download the app due to confusion about the purpose and utility of the app as well as its connection to the in-person curriculum, resulting in a limited use of the app outside of the in-person sessions. Additionally, some youth did not have access to a smartphone at all, which not only prevented them from accessing the app outside of the in-person sessions, but also contributed to the participants feeling left out of the program. A health educator described the experience of 1 youth who was homeless as such:

> *One participant mentioned that she felt like she was being discriminated against because she didn’t have a cell phone....* [Interview, Year 1]

Health educators also noted a lack of integration between the in-person elements of the curriculum and the ITK app. Because many of the ITK app features and activities were explained at the end of the modules, health educators commented on the difficulties of transitioning between the in-person curriculum and technology-based activities. One health educator stated the following:

> *I wish there was more involvement of the app in the actual curriculum… Like, yes, there is the whole, you know, app introduction for each module after their curriculum. But I wish it was something that we can use tied into our curriculum... It just kind of seems like the little, little side perk to the class—which it is, it is a perk, because like the other youth, who have not participated in the program, don't get to experience the app or get to have the information on the app. But I think it would still really help if we can actually use the app for facilitating, and that the youth can go back on the app and look through things that we've talked about already, or stuff like that.* [Interview, Year 3]

Youth Engagement

Overall, the health educators reported that youth were engaged and interested in the curriculum. They stated that participation and engagement increased among the youth when playing games with Kahoot, an online learning platform that allowed educators to gamify content delivery. One health educator explained it as such:

> *Oh, Kahoot. When it's working, it works great. Like when it's working, it's probably like the one thing that the youth get excited about, maybe because they already know what it is and they get to play it at school already. So they think right away, like, “Oh, yes, it's a game!”* [Interview, Year 3]

Youth also responded particularly well to activities utilizing the O*Net OnLine website, a free online career exploration tool. However, health educators also noted that youth preferred participatory activities in general, whether technology-based activities and games or in-person activities such as role plays compared to lecture-based activities. One health educator described their experience as follows:

> *Sometimes we're not using the tablets or we're not doing like any kind of more of a group discussion. Like when there's listening in or something, or when I'm just asking them questions, it's really hard to, it's like school. It's like, okay, raise your hand or something like that. That's where I start to lose them.* [Interview, Year 3]

Despite fostering interest and engagement, in some instances, the presence of electronic devices was distracting for some youth. One health educator described a common experience in an implementation log as follows:

> *Some youths had earphones plugged in the tablets, played games, or even took photos of themselves during the class time. Facilitators would walk around the room to ask the youth to stop playing with the tablets while a facilitator was presenting. Although facilitators had to tell the youth from time to time to stop being on the tablets, facilitators did their best to move the class along with fewer distractions.* [Log, Year 1]

Another common youth engagement issue was the need to contextualize or personalize content for the participants. On almost every implementation log, the health educators noted instances where they had to reframe content or add explanations. For example, 1 health educator noted their role in providing supplemental information regarding a video on the biology of conception and pregnancy:

> *Youth did not seem to understand the video as far as the feedback that we got after when trying to discuss. Facilitator replayed the video and broke it down into different wording with each section.* [Log, Year 3]

Discussion

Principal Findings

These findings illustrate some of the successes and challenges of integrating digital technology into an in-person sexual health education program from the critical perspective of health educators. As previous studies found, health educators commonly reported that technological issues such as connectivity and device compatibility affected implementation, which were not unique to sexual health education [9,17,18]. However, technological issues became less frequent over time, likely for 3 reasons. First, health educators gained experience and confidence in addressing common technological challenges, including making innovative adaptations or finding alternatives when technology malfunctioned. Second, additional training may have led to greater familiarity and comfort with the myriad of platforms and implementation strategies. Third, health educators provided ongoing feedback to the developers, which resulted in changes to certain technology features and problem resolution. The decline in technological issues demonstrates the importance of ongoing and iterative quality improvement processes and the need for sustained engagement by the app development team in any technology-based health education intervention. It also illustrates the need to ensure that health educators are comfortable and confident in using technology, either through prior experience or through training.

Despite the implementation challenges, the health educators held positive views about the value that technology added to the in-person education, particularly in engaging youth with the material. Overall, youth tended to be more involved when they actively interacted with the content, whether through the technology-based components or in-person activities. Technology may be one of many tools that can increase the interactivity of curricular content [12]. A review of a variety of computer-based technologies found that digital games had the most evidence supporting their use to increase student engagement [30]. Game-based activities were successful, supporting the evidence that well-designed gamification can increase student engagement and motivation, and demonstrating the potential for gamification of educational content [31,32]. While the digital content was generally well received by youth, health educators also noted that the technology-based activities were not fully integrated into the curriculum. This was similar to the findings by another study of an online sexual health education course where some teachers reported difficulties transitioning between web-based and in-person activities [9].

Although adolescents have widely adopted technology, our findings are reflective of research showing ongoing disparities in technology access and use at the individual, community, and institutional level [7,33]. While ITK was designed for youth in underserved settings including foster care and shelters, health educators were more likely to encounter technology issues such as lack of Wi-Fi and other hardware in non–school settings. This made the implementation of the technology components of the program more challenging [34]. Additionally, while most participants had phones, some had limited storage or shared the phone with other family members, making them less inclined to download or keep an app, particularly one that stored sensitive information. By contrast, other studies have found that youth appreciate the anonymity available through technology-based sexual health interventions [34].

While technology can enhance youth engagement and comprehension, this study highlighted the critical role of health educators who secure the hardware necessary for implementation, adapt the curriculum when technology fails, and contextualize and personalize digital content to meet the unique needs of the youth they serve. Other studies have demonstrated the importance of staff training, confidence, and self-efficacy for the success of efforts to integrate mobile technology into education [35,36]. Beyond technological competence, health educators also need the core capabilities in knowledge and skills to deliver effective, inclusive, and appropriate sexual health education, particularly when discussing sensitive sexual and reproductive health topics [37].

Limitations

A few limitations should be noted. The implementation log data is self-reported, so health educators may have underreported issues or interpreted a situation differently. However, these results also were consistent with annual interviews with the health educators. This study did not assess the prior experience or comfort level of the health educators with technology. In addition, because the ITK app changed over time in response to feedback and updates, some of the technical components or issues may have been resolved over time or varied by time period.

Conclusion

As more sexual health educational programs incorporate technology, they should consider the specific role and use of technological components from both a pedagogical and logistical standpoint. Developers should engage with youth and health educators when designing health curricula and apps to ensure that the content is integrated and promotes youth learning and engagement. App developers need to invest in usability testing and a system for reporting issues throughout implementation and iteratively update the product based on that feedback. Similarly, developers and organizations need to ensure that health educators have the training, confidence, and support necessary for successful implementation, including the curricular content, classroom management skills, and necessary technology.

Although technology is often presented as a solution to reach underserved populations, that premise is not yet fully realized. Educational programs considering the adoption or integration of technology should assess the potential needs and technological capacity of the participants and settings.

Acknowledgments

The authors thank our partners at the Youth+Tech+Health initiative of ETR, the staff at Fresno Economic Opportunities Commission, and all the study sites, as well as youth who participated in this study. This study was funded by the Department of Health and Human Services, Family and Youth Services Bureau, Personal Responsibility Education Program Innovative Strategies, grant 90AP2688-01-00. The funder had no role in the data collection, analysis, interpretation, or publication of this manuscript or future results.

Authors' Contributions

MJD is the principal investigator of the study and is one of the primary authors of the manuscript. RT, SH, and MP coded and analyzed implementation logs and interviews. RT wrote the first draft of the introduction and results and edited the manuscript. MP provided training and oversight for fidelity monitoring and conducted health educator interviews. AGG provided support to this study's evaluation activities, including health educator interviews. JY provided support to this study's evaluation design. All authors reviewed and edited the manuscript.

Conflicts of Interest

None declared.

References

1. Sharples M, Arnedillo-Sanchez I, Milrad M, Vavoula G. Mobile Learning Small Devices, Big Issues. In: Technology-Enhanced Learning: Principles and Products. Dordrecht, Germany: Springer; 2009:233-249.
2. Anderson M, Jiang J. Pew Research Center. Teens, social media and technology 2018. 2018. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2020-01-22]
3. Giovanelli A, Ozer EM, Dahl RE. Leveraging technology to improve health in adolescence: a developmental science perspective. J Adolesc Health 2020 Aug;67(2S):S7-S13 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.02.020] [Medline: 32718517]
4. Pound P, Langford R, Campbell R. What do young people think about their school-based sex and relationship education? A qualitative synthesis of young people's views and experiences. BMJ Open 2016 Sep 13;6(9):e011329 [FREE Full text] [doi: 10.1136/bmjopen-2016-011329] [Medline: 27625058]
5. Brayboy LM, McCoy K, Thamotharan S, Zhu E, Gil G, Houck C. The use of technology in the sexual health education especially among minority adolescent girls in the United States. Curr Opin Obstet Gynecol 2018 Oct;30(5):305-309 [FREE Full text] [doi: 10.1097/GCO.0000000000000485] [Medline: 30153129]
6. Sheoran B, Silva CL, Lykens JE, Gamedze L, Williams S, Ford JV, et al. YTH StreetConnect: development and usability of a mobile app for homeless and unstably housed youth. JMIR mHealth uHealth 2016 Jul 14;4(3):e82 [FREE Full text] [doi: 10.2196/mhealth.5168] [Medline: 27417882]
7. Gao N, Hayes J. California's digital divide. Public Policy Institute of California. 2021 Feb. URL: https://www.ppic.org/wp-content/uploads/jtf-californias-digital-divide.pdf [accessed 2021-09-13]
8. Rideout V, Robb MB. The common sense census: media use by tweens and teens. Common Sense Media. San Francisco, CA; 2019. URL: https://www.commonsensemedia.org/sites/default/files/uploads/research/2019-census-8-to-18-full-report-updated.pdf [accessed 2021-09-12]
9. van Lieshout S, Mevissen F, de Waal E, Kok G. Long Live Love+: evaluation of the implementation of an online school-based sexuality education program in the Netherlands. Health Educ Res 2017 Jun 01;32(3):244-257 [FREE Full text] [doi: 10.1093/her/cyx041] [Medline: 28481977]
10. Lim MSC, Hocking JS, Aitken CK, Fairley CK, Jordan L, Lewis JA, et al. Impact of text and email messaging on the sexual health of young people: a randomised controlled trial. J Epidemiol Community Health 2012 Jan;66(1):69-74. [doi: 10.1136/jech.2009.100396] [Medline: 21415232]
11. Mustanski B, Greene GJ, Ryan D, Whitton SW. Feasibility, acceptability, and initial efficacy of an online sexual health promotion program for LGBT youth: the Queer Sex Ed intervention. J Sex Res 2015;52(2):220-230. [doi: 10.1080/00224499.2013.867924] [Medline: 24588408]
12. Martin P, Cousin L, Gottot S, Bourmaud A, de La Rochebrochard E, Alberti C. Participatory interventions for sexual health promotion for adolescents and young adults on the internet: systematic review. J Med Internet Res 2020 Jul 31;22(7):e15378 [FREE Full text] [doi: 10.2196/15378] [Medline: 32735217]
13. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. Cochrane Database Syst Rev 2010 Sep 08(9):CD006483 [FREE Full text] [doi: 10.1002/14651858.CD006483.pub2] [Medline: 20824850]
14. Badawy SM, Kuhns LM. Texting and mobile phone app interventions for improving adherence to preventive behavior in adolescents: a systematic review. JMIR mHealth uHealth 2017 Apr 19;5(4):e50 [FREE Full text] [doi: 10.2196/mhealth.6837] [Medline: 28428157]

15. Widman L, Nesi J, Kamke K, Choukas-Bradley S, Stewart JL. Technology-based interventions to reduce sexually transmitted infections and unintended pregnancy among youth. J Adolesc Health 2018 Jun;62(6):651-660 [FREE Full text] [doi: 10.1016/j.jadohealth.2018.02.007] [Medline: 29784112]
16. Kalke KM, Ginossar T, Shah SFA, West AJ. Sex ed to go: a content analysis of comprehensive sexual education apps. Health Educ Behav 2018 Aug;45(4):581-590. [doi: 10.1177/1090198117749259] [Medline: 29284294]
17. Hur JW, Shannon D, Wolf S. An investigation of relationships between internal and external factors affecting technology integration in classrooms. J Digit Learn Teach Educ 2016 May 25;32(3):105-114. [doi: 10.1080/21532974.2016.1169959]
18. Hew KF, Brush T. Integrating technology into K-12 teaching and learning: current knowledge gaps and recommendations for future research. Education Tech Research Dev 2006 Dec 5;55(3):223-252. [doi: 10.1007/s11423-006-9022-5]
19. Merrill KG, Merrill JC, Hershow RB, Barkley C, Rakosa B, DeCelles J, et al. Linking at-risk South African girls to sexual violence and reproductive health services: a mixed-methods assessment of a soccer-based HIV prevention program and pilot SMS campaign. Eval Program Plann 2018 Oct;70:12-24 [FREE Full text] [doi: 10.1016/j.evalprogplan.2018.04.010] [Medline: 29890449]
20. Decker MJ, Gutmann-Gonzalez A, Price M, Romero J, Sheoran B, Yarger J. Evaluating the effectiveness of an intervention integrating technology and in-person sexual health education for adolescents (In the Know): protocol for a cluster randomized controlled trial. JMIR Res Protoc 2020 Aug 07;9(8):e18060 [FREE Full text] [doi: 10.2196/18060] [Medline: 32763885]
21. IDEO. The field guide to human-centered design. Design Kit. 2015. URL: https://www.designkit.org/resources/1 [accessed 2020-01-27]
22. Fakoya I, Cole C, Larkin C, Punton M, Brown E, Ballonoff Suleiman A. Enhancing human-centered design with youth-led participatory action research approaches for adolescent sexual and reproductive health programming. Health Promot Pract 2021 Apr 15;23(1):25-31. [doi: 10.1177/15248399211003544] [Medline: 33858267]
23. Gavin LE, Catalano RF, David-Ferdon C, Gloppen KM, Markham CM. A review of positive youth development programs that promote adolescent sexual and reproductive health. J Adolesc Health 2010 Mar;46(3 Suppl):S75-S91. [doi: 10.1016/j.jadohealth.2009.11.215] [Medline: 20172462]
24. Damon W. What is positive youth development? Ann Am Acad Pol Soc Sci 2004 Jan 01;591(1):13-24. [doi: 10.1177/0002716203260092]
25. 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]
26. Dey I. Grounding Grounded Theory: Guidelines for Qualitative Inquiry. San Diego, CA, US: Academic Press; 1999.
27. Potter WJ, Levine-Donnerstein D. Rethinking validity and reliability in content analysis. J Appl Commun Res 2009 May 21;27(3):258-284. [doi: 10.1080/00909889909365539]
28. Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook. Los Angeles, CA, US: Sage Publications, Inc; 1994.
29. Dedoose Version 8.0.35. Web application for managing, analyzing, and presenting qualitative and mixed method research data. Dedoose. Los Angeles, CA: SocioCultural Research Consultants, LLC; 2018. URL: https://www.dedoose.com/ [accessed 2020-05-15]
30. Schindler LA, Burkholder GJ, Morad OA, Marsh C. Computer-based technology and student engagement: a critical review of the literature. Int J Educ Technol High Educ 2017 Oct 2;14(1):1-28 [FREE Full text] [doi: 10.1186/s41239-017-0063-0]
31. Dixson MD. Creating effective student engagement in online courses: what do students find engaging? JoSoTL 2010 Jun;10(2):1-13 [FREE Full text]
32. Ofosu-Ampong K. The shift to gamification in education: a review on dominant issues. J Educ Technol Syst 2020 Apr 21;49(1):113-137. [doi: 10.1177/0047239520917629]
33. Benda NC, Veinot TC, Sieck CJ, Ancker JS. Broadband Internet access is a social determinant of health!. Am J Public Health 2020 Aug;110(8):1123-1125. [doi: 10.2105/AJPH.2020.305784] [Medline: 32639914]
34. Teadt S, Burns JC, Montgomery TM, Darbes L. African American adolescents and young adults, new media, and sexual health: scoping review. JMIR mHealth uHealth 2020 Oct 05;8(10):e19459 [FREE Full text] [doi: 10.2196/19459] [Medline: 33016890]
35. Cheng S, Lu L, Xie K, Vongkulluksn VW. Understanding teacher technology integration from expectancy-value perspectives. Teach Teach Educ 2020 May;91:103062. [doi: 10.1016/j.tate.2020.103062]
36. Laura N, Corlett D. Reflections on success: A retrospective of the mLearn conference series 2002-2005. Banff, Canada: mLearn 2006: Across generations and cultures; 2007:1-29.
37. Core Competencies Subcommittee of the California Adolescent Sexual Health Work Group (ASHWG). Core competencies for adolescent sexual and reproductive health. California Prevention Training Center. 2008 Sep 01. URL: https://www.stdhivtraining.org/x/STD_PM_Online_2011/HO8-Additional_Resources.pdf [accessed 2021-10-20]

Abbreviations

ITK: In the Know

LGBTQ: Lesbian, gay, bisexual, transgender, and queer or questioning

Edited by A Kushniruk; submitted 18.06.21; peer-reviewed by I Shubina, M Salimi, P Martin; comments to author 27.08.21; revised version received 22.10.21; accepted 06.11.21; published 12.01.22.

Please cite as:
Decker MJ, Harrison S, Price M, Gutmann-Gonzalez A, Yarger J, Tenney R
Educators' Perspectives on Integrating Technology Into Sexual Health Education: Implementation Study
JMIR Hum Factors 2022;9(1):e31381
URL: https://humanfactors.jmir.org/2022/1/e31381
doi:10.2196/31381
PMID:35019842

Original Paper

Adoption of the Website and Mobile App of a Preventive Health Program Across Neighborhoods With Different Socioeconomic Conditions in the Netherlands: Longitudinal Study

Elena Agachi[1*], MSc; Tammo H A Bijmolt[1*], Prof Dr; Jochen O Mierau[2*], Prof Dr; Koert van Ittersum[1*], Prof Dr

[1]Department of Marketing, Faculty of Economics and Business, University of Groningen, Groningen, Netherlands

[2]Department of Economics, Econometrics & Finance, Faculty of Economics and Business, University of Groningen, Groningen, Netherlands

[*]all authors contributed equally

Corresponding Author:
Elena Agachi, MSc
Department of Marketing
Faculty of Economics and Business
University of Groningen
PO Box 800
Groningen, 9700 AV
Netherlands
Phone: 31 50 363 3686
Email: e.agachi@rug.nl

Abstract

Background: Socioeconomic disparities in the adoption of preventive health programs represent a well-known challenge, with programs delivered via the web serving as a potential solution. The preventive health program examined in this study is a large-scale, open-access web-based platform operating in the Netherlands, which aims to improve the health behaviors and wellness of its participants.

Objective: This study aims to examine the differences in the adoption of the website and mobile app of a web-based preventive health program across socioeconomic groups.

Methods: The 83,466 participants in this longitudinal, nonexperimental study were individuals who had signed up for the health program between July 2012 and September 2019. The rate of program adoption per delivery means was estimated using the Prentice, Williams, and Peterson Gap–Time model, with the measure of neighborhood socioeconomic status (NSES) used to distinguish between population segments with different socioeconomic characteristics. Registration to the health program was voluntary and free, and not within a controlled study setting, allowing the observation of the true rate of adoption.

Results: The estimation results indicate that program adoption across socioeconomic groups varies depending on the program's delivery means. For the website, higher NSES groups have a higher likelihood of program adoption compared with the lowest NSES group (hazard ratio 1.03, 95% CI 1.01-1.05). For the mobile app, the opposite holds: higher NSES groups have a lower likelihood of program adoption compared with the lowest NSES group (hazard ratio 0.94, 95% CI 0.91-0.97).

Conclusions: Promoting preventive health programs using mobile apps can help to increase program adoption among the lowest socioeconomic segments. Given the increasing use of mobile phones among disadvantaged population groups, structuring future health interventions to include mobile apps as means of delivery can support the stride toward diminishing health disparities.

(JMIR Hum Factors 2022;9(1):e32112) doi:10.2196/32112

KEYWORDS

eHealth; mHealth; mobile health; mobile app; internet; preventive health program; health disparities; NSES; program adoption; survival analysis

Introduction

Background

Noncommunicable diseases currently account for more than half of the global burden of disease, causing an ever-increasing proportion of premature deaths in both low- and high-income countries [1]. This occurrence is driven by preventable factors such as unhealthy diets, lack of physical activity, and tobacco and alcohol consumption [2,3]. Although noncommunicable diseases affect all segments of the population, the socially disadvantaged groups experience higher risk factors for these diseases [4,5] while remaining hard to reach through preventive health interventions [6].

Socioeconomic disparities in the adoption of preventive health programs are a well-known challenge [7,8], with programs delivered via the web being a potential solution. Web-based health programs show higher prospects in terms of behavior change ability and accessibility as compared with offline programs [9-11], having predominantly two means of delivery: website and mobile app. However, no clear understanding exists yet as to which (if any) of these delivery means is better able to promote increased inclusivity of all population segments [12,13], or on the contrary, leads to reinforcement, or even widening, of the existing disparities [14,15].

Objective

The health program examined in this study is the SamenGezond (from Dutch: Healthy Together) platform. Originally introduced by the health insurance company Menzis, the health program is aimed at the general Dutch population, following the goals of improving the health behaviors and wellness of its participants. The SamenGezond program was originally offered in the form of a website (introduced in 2012) and subsequently expanded to also include a mobile app (in 2017). Currently, reaching approximately 1 million participants, the program offers a set of activities and coaching that support healthy nutrition, physical activity, and other health behaviors.

With an increasing proportion of health programs being delivered through the web [16], it is of added value to gain a better understanding of the potential differential impact that the delivery means can have on program adoption. Therefore, the objective of this study is to analyze whether the adoption of a preventive health program by especially the low socioeconomic segment differs between the website and the mobile app. Identifying whether either of these means of delivery can achieve a better adoption rate among the socially disadvantaged groups could allow for future refining of health policy tools, contributing toward alleviating existing health disparities.

Methods

Study Sample

The database analyzed in this study originated from the SamenGezond platform, a large-scale web-based health program delivered by a health insurance company to the general Dutch population. The health program was introduced in July 2012, initially through a website, and starting in October 2017, expanded to also include a mobile app. The website of the health program mainly comprises information and coaching related to wellness and healthy lifestyle. The mobile app is an extension of the website, introducing several additional features to the health program, such as the ability to record activities with GPS, interact with an internet-based coach, and set and complete health goals. All program participants initially used the website of the program, with a subsequent choice of enrolling for the mobile app or continuing the use of solely the website. Although the health program is offered by a health insurance company, its participants are not solely clients of this company but can also be insured elsewhere (health insurance participation being legally mandated in the Netherlands).

The health program's aim is to improve the lifestyle and wellness of its participants by focusing on physical activity, healthy eating habits, social activity, mental health, good sleep habits, and minimized stress. The activities that are provided involve entering or recording physical activities, reading articles, setting goals, including friends in challenges, answering health questions, being assisted by an internet-based coach, and forming a daily *fit-score* based on the individual activities within the platform. The program also offers benefits in the form of accumulated points from participation in the various sections of the platform, which can be used to acquire specific products, vouchers for various services, gadgets, or charity contributions.

Enrollment to the health program was open and free, and all the participants involved in this study provided their voluntary and informed consent. Approval for this project was obtained from the institutional review board of the University of Groningen.

Data were collected between 2012 and 2019 and analyzed in 2020 and 2021 within a longitudinal, nonexperimental study design. This study design was used as it allows for the examination of the duration until adoption of the 2 components of the health program for a large group of participants. The analyzed data had a weekly frequency, covering 376 weeks and including 83,466 participants. All program participants were aged >18 years and were residents of the Netherlands; no additional exclusion criteria were applied. When selecting the participants for this study, out of the 838,500 individuals who enrolled in the health program at the time, 404,398 (48.23%) individuals who had logged into the health program at least once in the past 2 years were examined for eligibility. Owing to limitations in data transfer and storage, approximately 24.73% (100,000/404,398) of the eligible group were invited randomly to participate in this study, with 83.46% (83,466/100,000) of them having provided their consent for participation. It is not possible to compare the analyzed sample with the approached sample, as no data were available on the individuals who did not provide their consent for data sharing.

Measures

The effectiveness of health programs is defined by their ability to contribute to disease prevention, which critically hinges on individuals adopting and using the program. However, a reason causing overall ambiguity related to the benefits of web-based health programs is the significant number of programs that have been unsuccessfully adopted by individuals [17]. In addition, a slow rate of adoption can serve as an early indicator of potential

dropout [18] and, subsequently, program failure, making it paramount to gain a better understanding of the health program adoption process, to support program success.

Building on the diffusion of innovations literature [19], this study measured the adoption of the health program using the number of individuals who signed up for the program each week (weekly subscription rate), with the rate of adoption being defined as the speed at which the health program spreads among the target group. Given that the adoption decision of individuals varies between technologies [20], this study focused in particular on the comparison of the rate of adoption of the health program between the website and mobile app across socioeconomic groups.

Examining the rate of adoption by distinguishing between population groups allows for assessing whether there are differences in the reach of the health program depending on the means of delivery. Given that solely individual factors offer insufficient explanations of differences in health behaviors [21], the neighborhoods in which individuals live have emerged as contexts affecting both health behaviors [22] and health outcomes [23]. On the basis of the discussion by Duncan and Kawachi [24], neighborhood socioeconomic status (NSES) was used in this study to distinguish between population segments with different socioeconomic characteristics.

The NSES measure was created in this study using data on key indicators for each neighborhood in the Netherlands, provided by the Central Bureau of Statistics [25]. Following the methodology outlined in the study by Dekker et al [26], the NSES measure was calculated using nonlinear iterative partial least squares principal component analysis on the following characteristics, given on a postcode level: average income, average property value, subsidized renting, share of high-income households, share of owner-occupied properties, share of low-income households, share of the population receiving unemployment benefits, share of the population receiving disability benefits, and share of the population receiving short-term unemployment benefits. NSES quintiles were used in the analysis of this study based on the constructed NSES measure, with a lower NSES quintile corresponding to lower levels of socioeconomic conditions.

To control for individual characteristics of the program participants, gender and age were included in the analysis as additional covariates. Moreover, as it can be expected that marketing campaigns that support the health program influence the rate of program adoption, the analysis was augmented with indicators for marketing activities taking place in each observed week or in the preceding week (to control for a lagged impact of marketing). The marketing campaigns considered were radio, television, and web-based campaigns.

Statistical Analysis

This study used survival modeling, which encompasses statistical procedures aimed at analyzing the time until an event occurs; the event of interest in this study was the adoption of the health program. The baseline population of individuals who could decide to adopt the health program was the general Dutch population, which was the target group of the health program (the marketing activities related to the health program took place in the Netherlands and were aimed at the Dutch population). Registration for the analyzed program was voluntary and free, which made it possible to examine the true rate of adoption, as in such a setup, individuals are free to decide by themselves when to enroll in the health program [27].

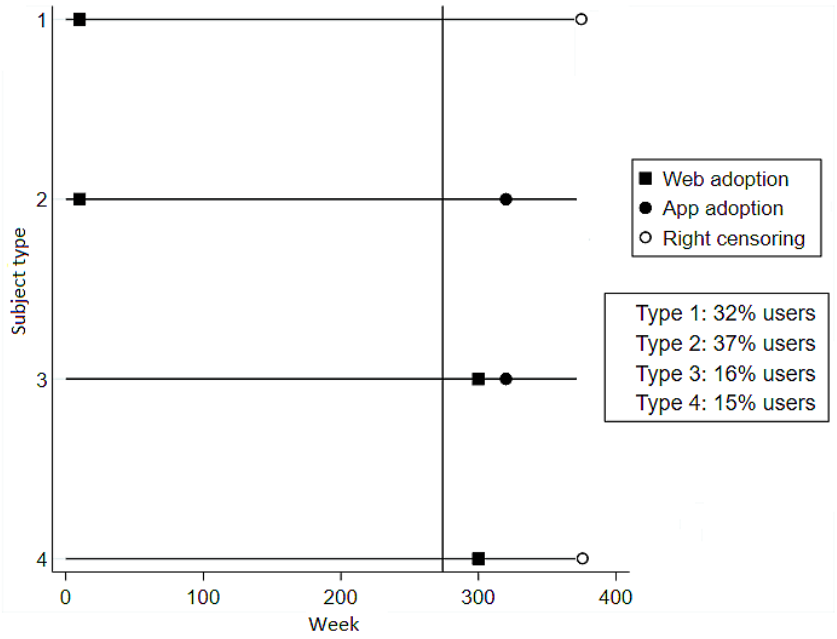
All the observed participants of the health program could experience two events: the adoption of the health program through the website and the adoption of the mobile app, with the former always preceding the latter. This pattern is schematically reflected in Figure 1, which depicts the several types of health program participants possible, depending on whether and when a participant adopted the website and the mobile app of the health program.

Figure 1 summarizes the 4 types of health program participants. Type 1 includes users (26,908/83,466, 32.24% of the program users) who adopted the health program before the mobile app was introduced (in the period between 2012 and 2017) and subsequently did not adopt the app when it became available in October 2017. Type 2 includes the users (31,333/83,466, 37.54%) who adopted the program before the app introduction and subsequently also adopted the app. Type 3 includes the users (12,979/83,466, 15.55%) who adopted the program after the app introduction and subsequently also adopted the app. Type 4 includes the users (12,246/83,466, 14.67%) who adopted the program after the app introduction and subsequently did not adopt the app. As the data set included solely participants who adopted the health program before September 2019, the observations linked to participants who would adopt the program or its mobile app after 2019 are not available and were censored.

At the end of the observation period, of the analyzed participants, 46.91% (39,154/83,466) had not adopted the app (yet) and were using only the website to access the health program. Possible reasons for this occurrence could be unawareness about the existence of the app or unwillingness to use the app based on not needing the extended features offered by it. Alternatively, the reluctance toward the adoption of the mobile app can be linked to privacy concerns [28].

As reflected in Figure 1, the adoption of the website and the mobile app were sequential events, with each participant being at risk for only 1 of these events at a time. The modeling approach that handles this structure best is an extension of the classical Cox model [29], namely the variants of the Prentice, Williams, and Peterson (PWP) model [30-32]. To answer the research question of this study related to differences in the rate of adoption between means of delivery, the PWP Gap–Time (PWP-GT) model is most appropriate, which estimates the effects of the following event since the time from the previous event [30,32]. This is achieved using time-dependent strata, where the hazard function is allowed to vary from event to event [33]. The PWP model estimates unbiased effects [33] and provides SEs robust to within-subject correlation [34]. Statistical analysis was performed using the survival package [35,36] implemented within the R (R Foundation for Statistical Computing) environment for statistical computing [37].

Figure 1. Types of health program participants.



Several robustness checks were performed for changes in the model specifications based on models accounting for nonproportional hazard rates, with the estimated parameters of the main PWP-GT model maintaining their direction and statistical significance (section 2 in Multimedia Appendix 1 [29,38,39], available on the web). Similarly, the models' parameter estimates remained unchanged when estimating the PWP-GT model with solely the NSES quintiles as covariates when controlling for being insured at the company that had initially introduced the health program and when accounting for subsequent program use measured by the number of weekly log-ins (section 3 in Multimedia Appendix 1, available on the web). Additionally, following the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) recommendations, the checklist presented in Multimedia Appendix 2 has been completed.

Results

Characteristics of Participants

On average, the 83,466 participants who were analyzed were in the program for >3 years (mean number of weeks in the health program 186, SD 124 weeks) and aged between 18 and 80 (mean age 46.5, SD 15.5) years, and 56% (46,741/83,466) of them were female. Table 1 provides an overview of the study participants' characteristics.

Among the participants' characteristics outlined in Table 1 is their distribution across NSES quintiles, which showed a higher proportion of program participants in the lowest 2 NSES quintiles. The insurance company operates on a larger scale in areas with low socioeconomic conditions, and as most participants were clients of this insurance company, the overrepresentation of the lowest NSES quintiles was reflected in the participants' distribution (a more detailed distribution of participants across NSES quintiles is discussed in section 1 in Multimedia Appendix 1, available on the web).

Table 1. Characteristics of study participants (N=83,466).

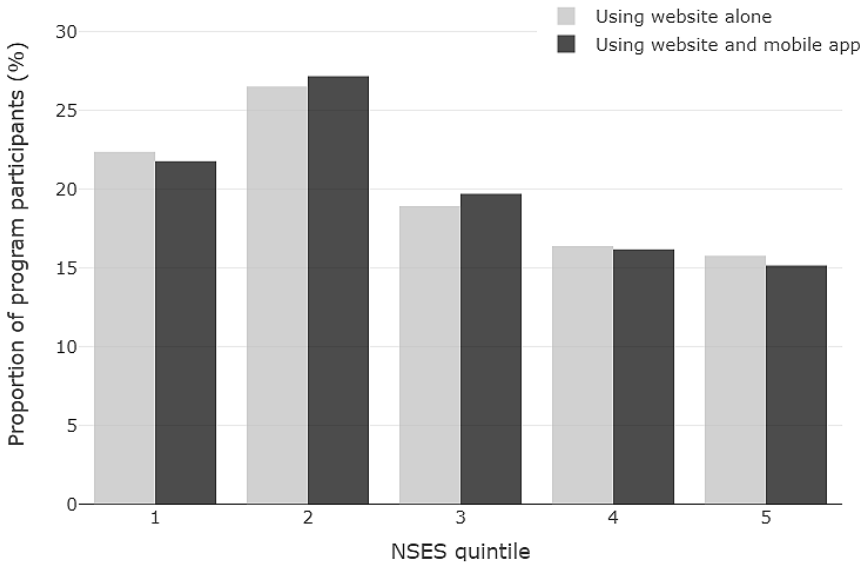
Key attributes	Values
Weeks covered, n	376
Number weeks in health program, mean (SD)	186 (124)
Participants using website alone, n (%)	39,146 (46.9)
Participants using website and mobile app, n (%)	44,320 (53.09)
Age (years), mean (SD)	46.5 (15.5)
Female participants, n (%)	46,741 (56)
Weeks with active marketing campaigns, n (%)	56 (14.9)
Participants per age group (years), n (%)	
18-26	8263 (9.89)
27-36	17,779 (21.3)
37-46	16,693 (19.99)
47-56	17,695 (21.2)
57-66	12,937 (15.49)
67-80	10,099 (12.09)
Participants per NSES[a] quintile (from lowest to highest socioeconomic conditions), n (%)	
First	18,446 (22.1)
Second	22,453 (26.9)
Third	16,109 (19.29)
Fourth	13,605 (16.3)
Fifth	12,853 (15.38)

[a]NSES: neighborhood socioeconomic status.

Figure 2 further depicts the proportion of program participants who were in each NSES quintile, separated into two groups: those who used solely the website of the health program and those who used both the website and the mobile app.

On the basis of Figure 2, the distribution of program participants across NSES quintiles showed a similar pattern independent of the health program's delivery means used.

Figure 2. Distribution of health program participants across neighborhood socioeconomic status (NSES) quintiles.



Regression Results

Analyzing the rate of adoption of the health program generated the results shown in Table 2, which contains the parameter estimates from the PWP-GT model accounting for NSES quintiles and the covariates of age, gender, and marketing indicators. The estimates shown in Table 2 are the hazard ratios (HRs; exponentiated model parameters) reflecting the effect size of the covariates, their corresponding 95% CIs, and *P* values.

Table 2. The impact of covariates on the rate of program adoption (Prentice, Williams, and Peterson Gap–Time model estimation results[a,b]).

Variables	Program adoption through website		Mobile app adoption	
	HR[c,d] (95% CI)	*P* value	HR[d] (95% CI)	*P* value
NSES[e] quintile				
First	1.000[f]	N/A[g]	1.000[f]	N/A
Second	1.034 (1.015-1.054)	.002	0.940 (0.907-0.973)	.001
Third	1.029 (1.008-1.051)	.02	0.954 (0.918-0.990)	.02
Fourth	1.031 (1.009-1.053)	.02	0.950 (0.912-0.988)	.02
Fifth	1.020 (0.997-1.043)	.12	0.948 (0.910-0.987)	.02
Age (in years)	1.007 (1.006-1.007)	<.001	0.980 (0.979-0.981)	<.001
Gender				
Female	1.000	N/A	1.000	N/A
Male	1.074 (1.060-1.088)	<.001	0.821 (0.797-0.845)	<.001
Marketing				
No	1.000	N/A	1.000	N/A
Yes	0.378 (0.360-0.396)	<.001	17.007 (16.979-17.035)	<.001

[a]Interpreting the estimated hazard ratios, for example, the second neighborhood socioeconomic status (NSES) quintile had an increased likelihood of program adoption via the website by a factor of 1.034 (95% CI 1.015-1.054) as compared with the lowest NSES quintile, keeping other covariates constant (equivalent to a 3.4% increased likelihood of adoption). On the other hand, the likelihood of adoption of the mobile app when comparing the second NSES quintile with the first one shows a decreased likelihood of adoption for the second NSES quintile by a factor of 0.940 (95% CI 0.907-0.973) or 6%.

[b]Observations=166,932 (the 166,932 observations reflect the 83,466 participants as the model accounts for 2 events per participant); R^2=0.255; maximum possible R^2=1.000; Wald test (*df*)=56,343.96 (14); *P*<.001.

[c]HR: hazard ratio.

[d]An HR of 1.000 was assigned to the reference level for each categorical covariate.

[e]NSES: neighborhood socioeconomic status.

[f]For the HR of 1.000 there is no 95% CI reported, as this is not an estimated HR, but is the default value assigned to the reference level.

[g]N/A: not applicable (it is the reference level).

The estimation results indicate that the impact of association with an NSES quintile on the rate of adoption of the health program differs between the 2 means of delivery. For the health program adoption through the website (shown in the first column of Table 2), most NSES quintiles have a statistically significant higher likelihood of adoption compared with the lowest NSES quintile; for example, an individual associated with the second NSES quintile has an increased likelihood of adoption by a factor of 1.034 (95% CI 1.015-1.054) compared with the lowest NSES quintile, keeping other covariates constant. However, for the adoption of the mobile app of the health program, all NSES quintiles have a lower likelihood of adoption compared with the lowest NSES quintile (shown in the second column of Table 2); comparing the second NSES quintile to the lowest one reveals a lower likelihood of adoption for the mobile app by a factor of 0.940 (95% CI 0.907-0.973), keeping other covariates constant. In addition, the estimated decrease in the likelihood of mobile app adoption for the higher NSES quintiles compared with the lowest one is higher than the estimated increased likelihood of website adoption (the effect sizes vary between a decreased likelihood of 4.6%-6% for the mobile app adoption and an increased likelihood of approximately 3% for the website adoption).

Examining the additional covariates shown in Table 2 indicate that older individuals have a higher rate of adoption of the website (HR=1.007, 95% CI 1.006-1.007) but a lower rate of adoption of the mobile app (HR=0.980, 95% CI 0.979-0.981) as compared with younger individuals. In addition, men are faster adopters of the website (HR=1.074, 95% CI 1.060-1.088) but slower adopters of the mobile app (HR=0.821, 95% CI 0.797-0.845) as compared with women. Finally, the weeks in which the marketing campaigns took place showed an increased rate of adoption for the mobile app (HR=17.007, 95% CI 16.979-17.035) but a decreased rate of adoption for the website (HR=0.378, 95% CI 0.360-0.396). The latter effect can be explained by the fact that all marketing campaigns took place

after the introduction of the mobile app, when the website rate of adoption had already slowed down. Focusing on the impact of marketing campaigns on the rate of adoption of the mobile app based on interaction terms (section 3 in Multimedia Appendix 1, available on the web), it turns out that higher NSES quintiles are more sensitive to the marketing campaigns than the lowest NSES quintile.

Discussion

Principal Findings

Owing to the increasing need for the prevention of risk behaviors such as poor nutrition habits and insufficient physical activity, web-based health programs have emerged as sustainable means of providing large-scale preventive health services to the population. As lower socioeconomic segments frequently exhibit lower uptake levels of preventive health services, additional research is needed to identify whether the lower socioeconomic segments are more receptive to *web-based* health programs. In this study, we examine whether the website or mobile app delivery means of a web-based preventive health program can induce a higher likelihood of adoption among the population group with the lowest socioeconomic conditions.

Analyzing the distribution of health program participants across NSES quintiles revealed a higher proportion in the lowest 2 NSES quintiles. Although generally, higher socioeconomic segments tend to be more represented in preventive health programs [40], the overrepresentation of the lowest NSES quintiles observed in this study is linked to the particularities of the insurance company that introduced the program, which operates on a larger scale in areas with low socioeconomic conditions.

The main findings of this study show that the website of the health program is associated with a higher likelihood of adoption among the higher socioeconomic population groups (between 2% and 3% increased likelihood of adoption; *P* value between .12 and .002 depending on the NSES quintile), whereas the mobile app displays a higher likelihood of adoption among the lowest socioeconomic group (between 2% and 6% increased likelihood of adoption; *P* value between .02 and .001 depending on the NSES quintile). Additional findings originating from this study reveal that the individuals' demographic characteristics are also linked to differences in adoption per means of delivery, with younger women ($P<.001$) more likely to adopt the mobile app of the health program. Marketing campaigns are estimated to increase the likelihood of mobile app adoption: 170% ($P<.001$) increased likelihood of mobile app adoption during, or right after, the weeks in which the marketing campaigns about the health program took place.

Comparison With Previous Work

The findings in other existing research differ on the topic of health program adoption through mobile apps among socioeconomic groups. On the one hand, it is estimated that higher NSES segments are more likely to use health programs delivered through mobile apps [12,15] because of their possession of better digital skills [41] and easier access to technological devices [42]. On the other hand, when engaging with web-based health programs, individuals living in lower NSES areas do so mostly through mobile apps [13,43] while showing similar ease of use of mobile apps for health as that of groups with higher socioeconomic conditions [44,45].

For the findings in this study, a circumstance likely linked to the lower socioeconomic group showing a higher likelihood of mobile app adoption is the possession of digital skills. A characteristic of the Dutch population is the high levels of digital skills, with the Netherlands ranking highest in Europe on this scale [46]. An overall high level of digital proficiency removes the potential barriers that could prevent lower socioeconomic segments from engaging with health programs delivered through mobile apps. In addition, among communities with the lowest socioeconomic conditions in the Netherlands, there is a positive attitude toward web-based lifestyle programs [47]; this, combined with the high digital skills, potentially facilitates the adoption of the mobile app of the health program.

Overall, the topic of disparities in the use of preventive health services in the existing literature is supported by a general consensus that although the socially disadvantaged segments experience a heavier burden of behavioral risk factors and disease [48,49], they are generally the least represented group in preventive health services [40], an occurrence leading to an accelerating inverse social gradient [50]. With web-based health programs having the ability to achieve higher adoption and use rates [9-11], more research is warranted on the effects of specific delivery means of such programs on uptake, especially among the lower socioeconomic population groups.

The realization that mobile app delivery of preventive health programs can increase adoption among the lowest socioeconomic segment of the population has important implications for the future design of health programs. The current digital age is characterized by a higher prevalence of mobile phone use as compared with computer use, a pattern that is especially heightened in low socioeconomic groups [51,52], with mobile phone ownership and use also seeing a sharp increase in the low-income countries [53]. Given these tendencies, structuring future health programs to include a mobile app as a means of delivery can help to increase the adoption of such services among the disadvantaged socioeconomic segment, which can support the stride toward achieving health equity among all population groups.

In light of the growing health care expenditures and the associated health disparities, it is of importance for health insurance policies to encourage prevention over treatment. Given the higher burden of costs associated with the population segment with the lowest socioeconomic conditions, it is paramount to increase preventive health service use within this segment. Designing future health programs, including the use of mobile apps, can facilitate the increase in the use of such services by the lowest socioeconomic group, thus leading to cost savings and encouraging further investment toward large-scale, web-based prevention services.

Limitations and Future Research

This study's setup and analysis methods have several limitations.

First, as the adoption of the health program continued beyond the window of observation analyzed in this study, it is not confirmed that the effects observed here would maintain their validity when including the later adopters of the website and the mobile app. To verify whether the later adopters match the pattern discussed in this study, the current analysis can be replicated at a later stage of the health program's existence.

Second, as the data analyzed were retrospective, the analysis was limited to only a few covariates related to the study's participants. To overcome this restriction, the measure of NSES was used to reflect socioeconomic conditions, with the limitation of missing information on residential moves and not accounting for heterogeneity within neighborhoods. An extension of the current analysis would be to include a measure of individual socioeconomic status and compare the inferences based on the individual-level measure to the ones obtained here based on the neighborhood-level measure.

Third, the program analyzed has a particular structure of adoption, with all participants initially adopting the website and subsequently having the choice to adopt the mobile app. Such a structure can potentially lead to conservative estimates, as in this setup, the mobile app adopters are already aware of the program's existence, having engaged previously with the website. Overcoming this restriction can be achieved by allowing program participants to adopt solely the mobile app, this being a structure toward which the analyzed health program is currently migrating.

Fourth, the specific health program analyzed includes solely individuals aged >18 years and from a high-income country. Given that youths worldwide are increasingly using more web-based services, future analyses could include younger individuals and compare web-based health program adoption within a more heterogeneous sample of low- and high-income countries. Although the findings of this study are based on the participants of a specific health program, we believe that the inferences drawn can be applied to the contexts of other health technologies, mainly because of the size and diversity of the data analyzed.

Finally, it is important to realize that solely adopting a preventive health program does not contribute to improvements in health. Therefore, as a future extension of this study, the analysis can include the subsequent use of the health program, and its impact on health outcomes, while distinguishing between population groups with different socioeconomic conditions.

Conclusions

In this study, a large-scale web-based preventive health program promoted in the Netherlands was analyzed, focusing on its rate of adoption among socioeconomic groups. The mobile app of the health program was identified as a delivery means linked to a higher likelihood of program adoption among the population group with the lowest socioeconomic conditions. This finding suggests that future preventive health interventions can benefit from web-based delivery through mobile apps, especially in the light of the increasing use of mobile phones among the disadvantaged population segments.

Acknowledgments

The authors wish to thank Menzis and Dacadoo for providing the data necessary for this project. Funding for this project was provided by Menzis (the health insurance company that introduced the health program). The funder had no involvement in the study design, data analysis and interpretation, writing the report, or the decision to submit the report for publication.

Conflicts of Interest

EA is funded by Menzis (the health insurance company that introduced the health program) in her position as a PhD candidate.

Multimedia Appendix 1
Additional estimation results.
[DOC File , 221 KB - humanfactors_v9i1e32112_app1.doc]

Multimedia Appendix 2
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.
[PDF File (Adobe PDF File), 135 KB - humanfactors_v9i1e32112_app2.pdf]

References

1. Noncommunicable diseases. World Health Organization. URL: https://www.who.int/health-topics/noncommunicable-diseases [accessed 2021-03-14]
2. Benziger CP, Roth GA, Moran AE. The global burden of disease study and the preventable burden of NCD. Glob Heart 2016 Dec 01;11(4):393-397. [doi: 10.1016/j.gheart.2016.10.024] [Medline: 27938824]
3. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 08;388(10053):1659-1724 [FREE Full text] [doi: 10.1016/S0140-6736(16)31679-8] [Medline: 27733284]

4. Marmot M. Social determinants of health inequalities. Lancet 2005 Mar 25;365(9464):1099-1104. [doi: 10.1016/s0140-6736(05)74234-3]
5. de Boer WI, Buskens E, Koning RH, Mierau JO. Neighborhood socioeconomic status and health care costs: a population-wide study in the Netherlands. Am J Public Health 2019 Jun;109(6):927-933. [doi: 10.2105/ajph.2019.305035]
6. Nelson HD, Cantor A, Wagner J, Jungbauer R, Quiñones A, Stillman L, et al. Achieving health equity in preventive services: a systematic review for a national institutes of health pathways to prevention workshop. Ann Intern Med 2020 Jan 14;172(4):258. [doi: 10.7326/m19-3199]
7. Bretthauer M, Kalager M. Disparities in preventive health services: targeting minorities and majorities. Ann Intern Med 2020 Jan 14;172(4):287. [doi: 10.7326/m19-3899]
8. Broeders D, Das D, Jennissen R, Tiemeijer W, de Visser M. Van verschil naar potentieel: Een realistisch perspectief op de sociaaleconomische gezondheidsverschillen. WRR-Policy Brief 7. 2018. URL: https://www.wrr.nl/binaries/wrr/documenten/policy-briefs/2018/08/27/van-verschil-naar-potentieel.-een-realistisch-perspectief-op-de-sociaaleconomische-gezondheidsverschillen/PB+7_Sociaaleconomische+gezondheidsverschillen_web_2908.pdf [accessed 2022-01-24]
9. Dounavi K, Tsoumani O. Mobile health applications in weight management: a systematic literature review. Am J Prev Med 2019 Jun;56(6):894-903 [FREE Full text] [doi: 10.1016/j.amepre.2018.12.005] [Medline: 31003801]
10. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
11. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of web-based vs. non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
12. Bol N, Helberger N, Weert JC. Differences in mobile health app use: a source of new digital inequalities? Inform Society 2018 Apr 26;34(3):183-193. [doi: 10.1080/01972243.2018.1438550] [Medline: 26281194]
13. Graetz I, Huang J, Brand RJ, Hsu J, Yamin CK, Reed ME. Bridging the digital divide: mobile access to personal health records among patients with diabetes. Am J Manag Care 2018 Jan;24(1):43-48 [FREE Full text] [Medline: 29350505]
14. Robinson L, Cotten SR, Ono H, Quan-Haase A, Mesch G, Chen W, et al. Digital inequalities and why they matter. Inform Commun Society 2015 Mar 16;18(5):569-582. [doi: 10.1080/1369118X.2015.1012532]
15. Lim D, Norman R, Robinson S. Consumer preference to utilise a mobile health app: a stated preference experiment. PLoS One 2020 Feb 21;15(2):e0229546 [FREE Full text] [doi: 10.1371/journal.pone.0229546] [Medline: 32084250]
16. Digital health. World Health Organization. URL: https://www.who.int/health-topics/digital-health [accessed 2021-03-14]
17. Labrique A, Vasudevan L, Chang LW, Mehl G. H_pe for mHealth: more "y" or "o" on the horizon? Int J Med Inform 2013 May;82(5):467-469 [FREE Full text] [doi: 10.1016/j.ijmedinf.2012.11.016] [Medline: 23279850]
18. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
19. Rogers E. Diffusion of Innovations. 4th ed. New York: Simon and Schuster; 2010.
20. Sezgin E. Introduction to current and emerging mHealth technologies: adoption, implementation, and use. In: Sezgin E, Yildirim SO, Yildirim S, editors. Current and Emerging mHealth Technologies. Switzerland: Springer; 2018:1-6.
21. Berkman L, Kawachi I, Glymour M. Social Epidemiology. 2nd ed. Oxfordshire, United Kingdom: Oxford University Press; 2014:510-548.
22. Robinette JW, Charles ST, Almeida DM, Gruenewald TL. Neighborhood features and physiological risk: an examination of allostatic load. Health Place 2016 Sep;41:110-118 [FREE Full text] [doi: 10.1016/j.healthplace.2016.08.003] [Medline: 27583527]
23. Jiao J, Drewnowski A, Moudon AV, Aggarwal A, Oppert J, Charreire H, et al. The impact of area residential property values on self-rated health: a cross-sectional comparative study of Seattle and Paris. Prev Med Rep 2016 Dec;4:68-74 [FREE Full text] [doi: 10.1016/j.pmedr.2016.05.008] [Medline: 27413663]
24. Duncan D, Kawachi I. Neighborhoods and Health. 2nd ed. Oxford, UK: Oxford University Press; 2018.
25. Key indicators neighbourhoods Netherlands, 2016. Central Bureau of Statistics. URL: https://www.cbs.nl/nl-nl/maatwerk/2016/30/kerncijfers-wijken-en-buurten-2016 [accessed 2020-09-07]
26. Dekker L, Rijnks R, Mierau J. The health potential of neighborhoods: a population-wide study in the Netherlands. SSM Popul Health 2021 Sep;15:100867 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100867] [Medline: 34377761]
27. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. J Med Internet Res 2010 Feb;12(1):e3 [FREE Full text] [doi: 10.2196/jmir.1361] [Medline: 20147006]
28. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and facilitators of the use of mobile health apps from a security perspective: mixed-methods study. JMIR Mhealth Uhealth 2019 Apr 16;7(4):e11223 [FREE Full text] [doi: 10.2196/11223] [Medline: 30990458]
29. Cox D, Oakes D. Analysis of Survival Data. Boca Raton, Florida, United States: Chapman and Hall/CRC; 1984.
30. Amorim L, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. Int J Epidemiol 2015 Feb;44(1):324-333 [FREE Full text] [doi: 10.1093/ije/dyu222] [Medline: 25501468]

31. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika 1981;68(2):373-379. [doi: 10.1093/biomet/68.2.373]
32. Yadav CP, Lodha R, Kabra SK, Sreenivas V, Sinha A, Khan MA, et al. Comparison of statistical methods for recurrent event analysis using pediatrics asthma data. Pharm Stat 2020 Nov 02;19(6):803-813. [doi: 10.1002/pst.2032] [Medline: 32484295]
33. Castañeda J, Gerritse B. Appraisal of several methods to model time to multiple events per subject: modelling time to hospitalizations and death. Rev Colomb Estad 2010;33(1):43-61 [FREE Full text]
34. Kelly PJ, Lim LL. Survival analysis for recurrent event data: an application to childhood infectious diseases. Statist Med 2000 Jan 15;19(1):13-33. [doi: 10.1002/(sici)1097-0258(20000115)19:1<13::aid-sim279>3.0.co;2-5]
35. Therneau T. A package for survival analysis in R. R Package Version 3.2-13. 2021. URL: https://CRAN.R-project.org/package=survival [accessed 2022-01-18]
36. Therneau T, Grambsch P. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.
37. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.r-project.org [accessed 2020-09-07]
38. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69(1):239-241. [doi: 10.1093/biomet/69.1.239]
39. Therneau T, Crowson C, Atkinson E. Using time dependent covariates and time dependent coefficients in the cox model. The Survival Package. 2013. URL: https://cran.microsoft.com/snapshot/2018-07-06/web/packages/survival/vignettes/timedep.pdf [accessed 2021-02-01]
40. Jusot F, Or Z, Sirven N. Variations in preventive care utilisation in Europe. Eur J Ageing 2011 Oct 21;9(1):15-25. [doi: 10.1007/s10433-011-0201-9]
41. Angosto S, García-Fernández J, Valantine I, Grimaldi-Puyana M. The Intention to Use Fitness and Physical Activity Apps: A Systematic Review. Sustainability 2020 Aug 17;12(16):6641. [doi: 10.3390/su12166641]
42. Ha J, Kang SJ, Kim Y. Sport fans in a “smart sport” (SS) age: drivers of smartphone use for sport consumption. Int J Sports Market Sponsor 2017 Aug 07;18(3):281-297. [doi: 10.1108/ijsms-08-2017-093]
43. Chang E, Blondon K, Lyles CR, Jordan L, Ralston JD. Racial/ethnic variation in devices used to access patient portals. Am J Manag Care 2018 Jan 01;24(1):1-8 [FREE Full text] [Medline: 29350513]
44. Ramirez V, Johnson E, Gonzalez C, Ramirez V, Rubino B, Rossetti G. Assessing the use of mobile health technology by patients: an observational study in primary care clinics. JMIR Mhealth Uhealth 2016 Apr 19;4(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4928] [Medline: 27095507]
45. Miller DP, Weaver KE, Case LD, Babcock D, Lawler D, Denizard-Thompson N, et al. Usability of a novel mobile health iPad app by vulnerable populations. JMIR Mhealth Uhealth 2017 Apr 11;5(4):e43 [FREE Full text] [doi: 10.2196/mhealth.7268] [Medline: 28400354]
46. The Netherlands ranks among the EU top in digital skills. Central Bureau of Statistics. 2020. URL: https://www.cbs.nl/en-gb/news/2020/07/the-netherlands-ranks-among-the-eu-top-in-digital-skills [accessed 2021-05-20]
47. Faber JS, Al-Dhahir I, Reijnders T, Chavannes NH, Evers AW, Kraal JJ, et al. Attitudes toward health, healthcare, and eHealth of people with a low socioeconomic status: a community-based participatory approach. Front Digit Health 2021 Jul 8;3:690182 [FREE Full text] [doi: 10.3389/fdgth.2021.690182] [Medline: 34713165]
48. Kivimäki M, Batty GD, Pentti J, Shipley MJ, Sipilä PN, Nyberg ST, et al. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. Lancet Public Health 2020 Mar;5(3):e140-e149 [FREE Full text] [doi: 10.1016/S2468-2667(19)30248-8] [Medline: 32007134]
49. Allen L, Williams J, Townsend N, Mikkelsen B, Roberts N, Foster C, et al. Socioeconomic status and non-communicable disease behavioural risk factors in low-income and lower-middle-income countries: a systematic review. Lancet Global Health 2017 Mar;5(3):277-289. [doi: 10.1016/s2214-109x(17)30058-x]
50. Marmot M. Health equity in England: the Marmot review 10 years on. Br Med J 2020 Feb 24;368:m693. [doi: 10.1136/bmj.m693] [Medline: 32094110]
51. Mobile fact sheet. Pew Research Center, Internet & Technology. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/?menuItem=011fca0d-9756-4f48-b352-d58f343696bf [accessed 2021-05-20]
52. Marler W. Mobile phones and inequality: findings, trends, and future directions. New Media Society 2018 Apr 07;20(9):3498-3520. [doi: 10.1177/1461444818765154]
53. Silver L, Smith A, Johnson C, Jiang J, Anderson M, Rainie L. Use of smartphones and social media is common across most emerging economies. Pew Research Center, Internet & Technology. 2019. URL: https://www.pewresearch.org/internet/2019/03/07/use-of-smartphones-and-social-media-is-common-across-most-emerging-economies [accessed 2021-05-20]

Abbreviations

HR: hazard ratio
NSES: neighborhood socioeconomic status
PWP: Prentice, Williams, and Peterson

PWP-GT: Prentice, Williams, and Peterson Gap–Time
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by M Focsa; submitted 15.07.21; peer-reviewed by H Do, A Hidki, C Reis; comments to author 01.09.21; revised version received 13.10.21; accepted 05.11.21; published 02.02.22.

Please cite as:
Agachi E, Bijmolt THA, Mierau JO, van Ittersum K
Adoption of the Website and Mobile App of a Preventive Health Program Across Neighborhoods With Different Socioeconomic Conditions in the Netherlands: Longitudinal Study
JMIR Hum Factors 2022;9(1):e32112
URL: https://humanfactors.jmir.org/2022/1/e32112
doi:10.2196/32112
PMID:35107433

Original Paper

Developing an Educational Website for Women With Endometriosis-Associated Dyspareunia: Usability and Stigma Analysis

Abdul-Fatawu Abdulai[1], RN, MSc; A Fuchsia Howard[1], RN, PhD; Paul J Yong[2], MD, PhD; Heather Noga[3], MA; Gurkiran Parmar[2], BSc; Leanne M Currie[1], RN, PhD

[1]The School of Nursing, University of British Columbia, Vancouver, BC, Canada

[2]Department of Obstetrics & Gynaecology, University of British Columbia, Vancouver, BC, Canada

[3]Women Health Research Institute, British Columbia Women's Hospital & Health Center, Vancouver, BC, Canada

Corresponding Author:
Abdul-Fatawu Abdulai, RN, MSc
The School of Nursing
University of British Columbia
T201-2211 Wesbrook Mall
Vancouver, BC, V6T 2B5
Canada
Phone: 1 7782230137
Email: Fatawu@student.ubc.ca

Abstract

Background: Endometriosis is a chronic condition that affects approximately 10% of women worldwide. Despite its wide prevalence, knowledge of endometriosis symptoms, such as pelvic pain, and treatments remains relatively low. This not only leads to a trivialization of symptoms and delayed diagnosis but also fuels myths and misconceptions about pain symptoms. At the same time, the use of web-based platforms for information seeking is particularly common among people with conditions that are perceived as stigmatizing and difficult to discuss. The *Sex, Pain, and Endometriosis* website is an educational resource designed to provide evidence-based information on endometriosis and sexual pain to help people understand the condition, feel empowered, dispel myths, and destigmatize endometriosis-associated sexual pain.

Objective: The study objective is to evaluate the usability of the website and assess for destigmatizing properties of sexual health–related web-based resources.

Methods: We conducted a usability analysis by using a think-aloud observation, a postsystem usability questionnaire, and follow-up interviews with 12 women with endometriosis. The think-aloud data were analyzed using the framework by Kushniruk and Patel for analyzing usability video data, the questionnaire data were analyzed using descriptive statistics, and the follow-up interviews were analyzed using simple content analysis. We conducted a usability assessment by deductively analyzing the interview data via a trauma-informed care framework and a content analysis approach.

Results: Through usability analysis, we found the website to be simple, uncluttered, satisfying, and easy to use. However, 30 minor usability problems related to navigation; website response; the comprehension of graphics, icons, and tabs; the understanding of content; and mismatch between the website and users' expectations were reported. In our stigma analysis, we found the web content to be nonstigmatizing. The participants suggested ways in which websites could be designed to address stigma, including ensuring privacy, anonymity, inclusiveness, and factual and nonjudgmental content, as well as providing opportunities for web-based engagement.

Conclusions: Overall, the participants found the website to be useful, easy to use, and satisfying. The usability problems identified were largely minor and informed the website redesign process. In the context of the limited literature on stigma and website design, this paper offers useful strategies on how sexual health–related websites can be designed to be acceptable and less stigmatizing to individuals with sensitive health issues.

(JMIR Hum Factors 2022;9(1):e31317) doi:10.2196/31317

KEYWORDS

endometriosis; sexual pain; dyspareunia; usability testing; think-aloud; stigma; web sites; digital health; informatics

Introduction

Background

Endometriosis is a chronic condition in which endometrial-like tissue is present outside the uterus, typically in the pelvic cavity [1]. Although the actual prevalence of endometriosis may be difficult to quantify, it is estimated that the disease affects approximately 1 out of every 10 women worldwide [2,3]. Women with this disorder tend to experience a variety of symptoms, including sexual pain, menstrual pain, chronic pelvic pain, and infertility [4]. Although pelvic pain is the most common symptom of endometriosis, 50% of women also experience endometriosis-associated dyspareunia—pain experienced during or after penetrative vaginal intercourse [5,6]. Endometriosis-associated dyspareunia affects multiple aspects of life, leading to absenteeism from work, poor interpersonal relationships, and impaired social functioning, and has a profound negative impact on quality of life [7].

Despite the negative effect of sexual pain on the life of people with endometriosis, many patients and health care providers do not take this symptom seriously, whereas others neglect pain symptoms or are unaware of the link between painful sex and endometriosis [8]. Even if the health care provider and patient are both knowledgeable about dyspareunia, the private nature of sexual pain may inhibit patient–provider discussions on the topic [9]. Owing to fear of social stigma, some patients with dyspareunia may also feel reluctant to visit or disclose their sexual pain experiences to health care providers [9]. Indeed, stigma has been identified as a significant barrier to the uptake of sexual health–related interventions, including management of endometriosis-associated dyspareunia [7].

Given the potentially stigmatizing nature of sexual pain, websites have been identified as complementary tools for disseminating patient information concerning endometriosis and its related symptoms [10]. With the internet becoming a major source of health information, people can explore sensitive and intimate health topics such as sexual pain in a private setting. Available evidence suggests that the use of web-based platforms for information seeking is particularly common among people with conditions or symptoms that are perceived as stigmatizing and difficult to discuss [11,12]. Web-based information is particularly important as it may help patients understand the complex relationship between endometriosis-associated pelvic pain and painful sex and improve patients' access and adherence to recommended treatments while maintaining anonymity [8]. The advantages associated with access to web-based information on sensitive health topics present an opportunity to develop a web-based educational resource for people with endometriosis-associated dyspareunia. In response to this need, we developed a patient-centered educational website to provide evidence-based information on endometriosis-associated dyspareunia, which became the *Sex, Pain, and Endometriosis* website.

Sex, Pain, and Endometriosis Website Development

The website was developed by a multidisciplinary team of scientists, health care professionals, patient partners, and community organizations in close collaboration with a web design company as part of an end-of-grant knowledge translation project. We adopted a patient-oriented research approach throughout the website development process wherein patient partners were equal team members in recognition that patients provide critical experience-based perspectives that are essential for creating a meaningful website [13]. The research group drew on the Knowledge to Action Framework and technology-enabled knowledge translation for developing the website [14,15]. We first conducted a needs assessment with our patient partners to determine the scope of the website, content, esthetics, and key messages. Key findings of the needs assessment included prioritizing necessary information on the causes of endometriosis-associated sexual pain and options for treatment. Our patient partners also expressed their desire to develop a website that would promote inclusiveness (eg, diverse gender identities, sexual orientations, ethnocultural backgrounds, and ages) and help address the stigma of endometriosis-associated sexual pain. Although the idea of addressing stigma was not systematically thought through at the time of developing the website, the development team was aware of the need for the website to address stigma in one way or another. Therefore, the images and language used on the website were carefully designed so as not to offend or stereotype website users. We also conducted a landscape analysis of pre-existing endometriosis and sexual pain websites to determine their content and features. This was followed by an iterative product development. Following the release of version 1 of the website, we conducted the usability analysis reported in this paper to determine its usability and functionality and whether the website met the users' needs. The website was modified based on the findings of the usability analysis before the final launch in February 2021. The main purpose of this website is to help people understand endometriosis-associated dyspareunia, feel empowered, and dispel the myths and misconceptions surrounding endometriosis and sexual pain [9]. The website has six main sections: providing information on endometriosis, painful sex, causes of painful sex, and treatment for painful sex; resources; and frequently asked questions. The process of website development is described in a forthcoming publication. Figure 1 shows the homepage with the main sections of the website and Figure 2 shows infographics of people affected by endometriosis-associated sexual pains and other sections of the website.

The purpose of this study is 2-fold; it aimed to (1) evaluate the usability of the website and (2) assess for destigmatizing properties of sexual health–related web-based resources. The usability evaluation is expected to help improve the web design to make it easy to use, satisfying, and acceptable to end users. We assessed for destigmatizing properties as websites on sensitive health topics are not typically assessed for their ability to address or exacerbate stigma despite stigma being an outcome

of interest for content developers [16-18]. That is, it is unclear how the design of digital platforms for general sensitive health problems could help address stigma or inadvertently reproduce and perpetuate stigma among users [19,20].

Figure 1. Landing page of the website.

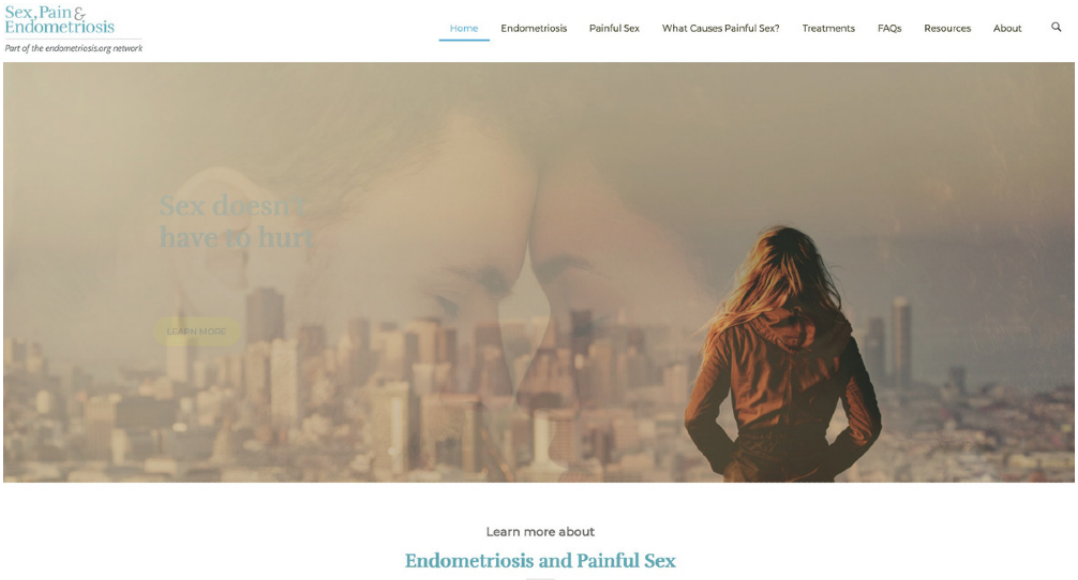
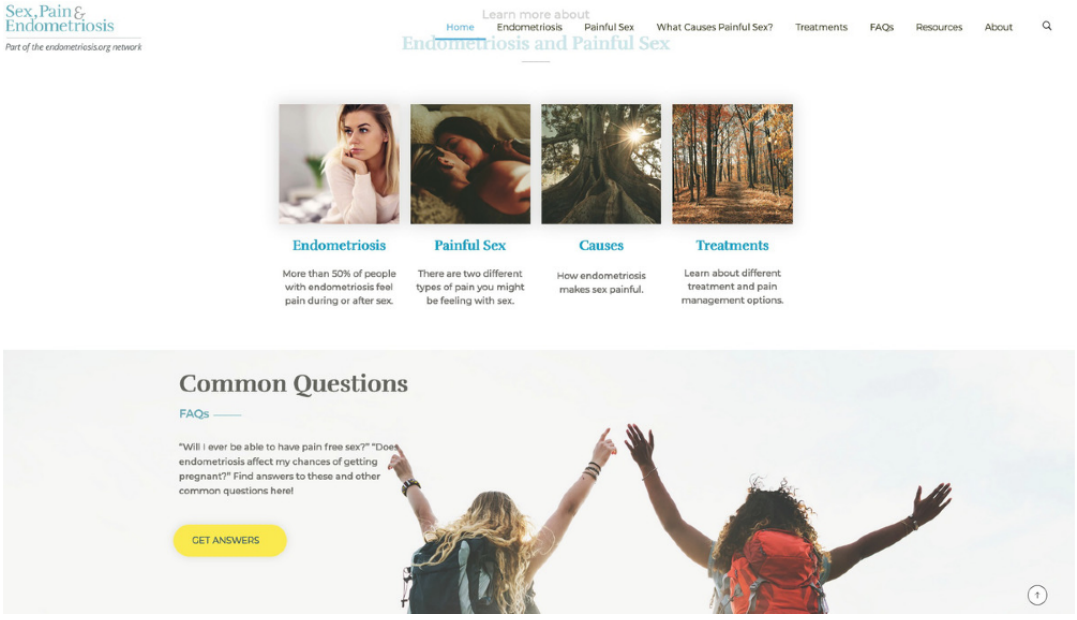


Figure 2. Other sections of the website.



Methods

Overview

We conducted a usability analysis by using think-aloud observations followed by post–think-aloud interviews and a postinterview questionnaire with 12 study participants. Using our website as a reference point, we also asked the participants how sexual health–related websites may be designed to address stigma. Owing to the COVID-19 pandemic, all activities were carried out on the web.

Recruitment

There have been no studies to our knowledge that provide population-based information on women with endometriosis and dyspareunia in western Canada. However, at the time of this study, there were approximately 300 patients diagnosed with endometriosis who were registered in the Endometriosis Pelvic Pain Interdisciplinary Cohort Data Registry of a large urban health care center in western Canada and who consented to be contacted for future research. Using a systematic sampling approach, we selected every 11th person from a list of approximately 300 patients in the data registry. A total of 45 participants were subsequently contacted. Participants who expressed interest but did not meet the inclusion criteria were excluded. The inclusion criteria were (1) patients aged ≥18 years; (2) new patients or patients rereferred to the center between May 1, 2019, and December 31, 2019; (3) consent to be contacted for future research; (4) patients known to have

clinically suspected or diagnosed endometriosis; and (5) experience of self-reported deep sexual pain (alone or partnered). The exclusion criteria were (1) not sexually active (alone or partnered), (2) never experienced sexual pain (alone or partnered), and (3) not fluent in English.

Ethics Approval

Ethics approval was obtained from the Children's and Women's Research Ethics Board at the University of British Columbia (research ethics board approval number: H19-03556). Of the 45 participants contacted, 12 (27%) consented to take part in the study.

Data Collection

Demographics

Before other data collection, the participants were asked to complete a demographic questionnaire that included age, gender identity, ethnicity, computer skills, and the frequency of computer use.

Usability Analysis

We conducted a usability analysis by using think-aloud observations, a Post-Study System Usability Questionnaire (PSSUQ), and a follow-up interview. In the think-aloud procedure, we asked the participants to carry out 5 task scenarios on our website while sharing their screens. The task scenarios represented different ways of searching for information about endometriosis-associated dyspareunia. These included (1) the meaning of *cul-de-sac* in relation to endometriosis, (2) different ways of treating sexual pain, (3) information on fertility, (4) anticipation of the pain cycle, and (5) the role of nervous system sensitization in painful sex. The participants were allowed to begin with any of the tasks. The think-aloud observations were conducted remotely via Zoom (Zoom Video Communications, Inc), all sessions were audio recorded, and the shared screen was video recorded. The recording did not include the participants' faces to protect their privacy. The PSSUQ was conducted to evaluate the overall system usefulness, information quality, and interface quality [21]. Adapted from Stinson et al [22], we conducted a follow-up interview to understand aspects of usability that were not made obvious during the think-aloud procedure.

Stigma Analysis

Given that some sexual health–related technologies may inadvertently exacerbate stigma [19,20], the follow-up interviews also included questions to understand the ways in which such websites can be designed to address stigma. In other words, the focus of the interviews was to use our website as a reference point to identify generalities about how destigmatizing sexual health–related websites could be designed. The interviews were audio recorded and transcribed verbatim by the first author (AFA).

Data Analysis

Usability Analysis

The think-aloud video data were analyzed by the first author based on the coding scheme by Kushniruk and Patel [23] for analyzing think-aloud data in patient information systems. Using the coding scheme as a guide, the first author watched and annotated the videotape and audio recordings with usability problems. Usability problems in the video and audio data were categorized into the five main thematic areas from the study by Kushniruk and Patel [23]: (1) navigation problems (related to the user finding desired information, icons, and labels), (2) comprehensiveness of graphics and text (problems related to the participants' understanding of labels, icons, and content), (3) system responses (how the website responded to the users' actions), (4) information content (which aspects of the system contained too much or too little information), and (5) mismatch between the website and the users' expectations. As the analysis proceeded, additional labels that were not captured in the initial coding scheme but emerged during the analysis were added. A second person (GP) independently reviewed all the video recordings, marked the usability problems, and categorized them based on the 5 thematic areas. The 2 reviewers' results were compared, and the differences were resolved by discussion. Content analysis procedures were used to identify additional usability problems in the follow-up interviews, whereas demographics and the PSSUQ were analyzed with descriptive statistics (1=minimum and 7=maximum, where lower scores indicate better satisfaction and higher scores indicate poor satisfaction). Problems identified from the follow-up interviews were also categorized under the 5 themes in the framework by Kushniruk and Patel [23].

Stigma Analysis

Data relating to potentially destigmatizing properties of sexual health–related websites were analyzed using a deductive approach to qualitative content analysis [24]. The deductive analysis followed a trauma-informed care framework [25] to identify the destigmatizing properties of websites. Trauma and stigma are inherently intertwined [26]. As such, we suggest that a trauma-informed care framework is relevant for analyzing data elicited from participants about stigma as it may inform recommendations to address stigma concerns among people who use sexual health–related websites [25,26]. The 5 principles of trauma-informed care by Fallot and Harris [25] were used to guide the analytic approach. These principles are (1) emotional safety (ensuring services are welcoming), (2) choice (ensuring individuals have options over their treatment and life), (3) collaboration (sharing power and making decisions with individuals), (4) trustworthiness (providing clear, credible, and consistent information about the condition), and (5) empowerment (providing an atmosphere that allows individuals to feel validated and affirmed). All transcripts were uploaded into NVivo software (version 11; QSR International). Using the steps for conducting deductive thematic content analysis by Braun and Clark [24], the first author initially familiarized himself with the data. Second, codes and concepts related to each of the 5 principles of the trauma-informed care framework were assigned to the text. Third, patterns and themes were searched for across the different interviews. Finally, the codes and concepts together with the subthemes were matched with their respective global themes, also known as the principles of trauma-informed care. This deductive approach allowed for the systematic identification of the participants' perspectives across 3 levels, including global themes (ie, the 5 principles of

trauma-informed care), subthemes, and concepts. Data saturation was achieved when no new concepts were identified in the data. The coding was discussed with coauthors FH and LC until a consensus was achieved.

Results

Demographic Information

The participants' ages ranged from 30 to 63 years, with a mean age of 38.75 (SD 8.55) years. All participants (12/12, 100%) self-identified as heterosexual, and all were from a large metropolitan area in western Canada. A total of 11 participants (11/12, 92%) self-identified as women, and 1 participant (1/12, 8%) did not disclose. Of the 12 participants, 6 (50%) identified as White, 2 (17%) identified as Hispanic, 1 (8%) identified as Indigenous, and 3 (25%) did not disclose their ethnic identity. All participants (12/12, 100%) reported using the internet daily, 42% (5/12) described their computer skills as very good, 50% (6/12) indicated that their computer skills were quite good, and 8% (1/12) indicated that their skills were neither good nor bad.

Task Completion

Tasks that required finding information in the *Mechanisms* section were the most difficult for the participants to complete, followed by finding the meaning of *cul-de-sac* in the *Endometriosis* section (Table 1). All participants (12/12, 100%) were able to locate information in the *Treatment* section, and 92% (11/12) of the participants located the information on fertility. Tasks were considered incomplete if the participants were assisted by the researcher (AF) or if they made several mistakes before locating the item. Table 1 shows the task completion rate and the average time (in seconds) it took the participants to complete each task. The entire study took approximately 1.2 hours for each participant to complete.

Table 1. Task completion rate (N=12).

Task number	Task	Website section	Participants completing the task, n (%)	Time (seconds), mean (SD)
1	Find the meaning of *cul-de-sac* in relation to endometriosis	Endometriosis	9 (75)	165 (23)
2	Find the different ways of treating sexual pain	Treatment	12 (100)	50 (15)
3	Find the information on fertility	FAQs[a]	11 (92)	80 (12)
4	Locate the anticipation of pain cycle	Mechanisms	8 (67)	206 (19)
5	Find the role of nervous system sensitization in painful sex	Mechanisms	7 (58)	260 (26)

[a]FAQs: frequently asked questions.

Overall Usability

Most participants expressed minimal difficulty with the think-aloud procedure, although a few had to be occasionally reminded to speak out loud. All participants (12/12, 100%) were able to complete the tasks, although 3 participants (3/12, 25%) needed some hints such as *please click on this link* or *the tab is located up there*. Generally, the participants were happy with the layout of the website. Participant 3, a woman aged 63 years, said the following:

> *This is a very simple website...I am not bombarded with too much information, the writing is good for my age...I don't have to strain my eyes to read this.*

Usability Problems

The think-aloud observations and the postsession interviews produced 30 usability problems, of which 23 (77%) were identified via the think-aloud process, and 7 (23%) were identified in the postsession interviews. Table 2 contains the problems identified from the usability analysis organized according to the 5 thematic areas by Kushniruk and Patel [23], the location of each problem on the website, the number of times the problem occurred, and the number of users encountering the problem. The problems were largely content-related issues, particularly related to a preponderance of provider perspectives on the website. For instance, it was generally agreed that the content of the website was medically oriented and currently lacked patients' perspectives on endometriosis and sexual pain. The participants also suggested solutions to some of the usability problems they identified (Table 3). The mean overall PSSUQ score was 2.41 (SD 0.85) in the range of 1 to 7, indicating relatively high satisfaction. Although the participants' scores across the 3 PSSUQ metrics were all below the average PSSUQ scores, system usefulness had the lowest score, indicating a better metric of usability. This was followed by interface quality and information quality. Table 4 shows the mean score for each PSSUQ item as well as the overall mean score for system usefulness, information quality, and interface quality.

Table 2. Interface problems from usability analysis.

Category and number	Usability problem	Page location	Times occurred, n	Users encountering the problem, n (%)
Navigation problems				
1	There is no search bar.	Home	15	12 (100)
2	The information underneath the homepage is not apparent to users.	Home	8	5 (42)
3	Links to other websites open on the same page. Difficult to navigate back to main page.	Home	6	3 (25)
4	The *go to top* icons at the bottom of the pages are not immediately visible.	Several	5	2 (17)
5	The sexual response cycle diagram does not fit in the screen for a whole view.	Mechanism	6	3 (25)
6	In-text references are not directly linked to the reference list. Users have to scroll up and down in search of references.	Several	5	5 (42)
7	Treatment pop-ups are too small. Not convenient to users.	Treatment	9	6 (50)
8	Not enough hyperlinks and hypertext to redirect users to different but related pages.	Several	7	5 (42)
Comprehension of graphics and text				
9	*Mechanism* section not clearly understood.	Home	14	12 (100)
10	The *slider* affordance on the diagram showing signs of endometriosis portrays a click function rather than a slider.	Endometriosis	6	8 (67)
11	Users did not understand the term *dyspareunia.*	Pain types	5	8 (67)
12	Not enough affordances to prompt users to click on diagrams and text.	Several	8	6 (50)
13	Too much content in treatment options and *mechanisms*.	Treatment	4	3 (25)
System response				
14	The link to *entry pain* does not respond.	Pain types	9	7 (58)
15	Pages load quite slowly. Takes an average of 8-10 seconds.	Several	8	7 (58)
16	Meaning on labels not immediately apparent to users.	Several	5	4 (33)
Information content				
17	Too much text in treatment pop-up.	Treatment	6	3 (25)
18	Endometriosis is not explained on the homepage.	Home	3	3 (25)
19	Not enough content on symptoms except the description of sexual pain.	Symptoms	2	2 (17)
20	Content on treatment pop-ups is too cluttered.	Treatment	5	4 (33)
21	Patient perspectives or voices are lacking on the website. It is medically oriented.	Several	3	4 (33)
Mismatch between the system and users' expectations				
22	Images are too cheerful to portray feelings of pain.	Symptoms	4	5 (42)
23	Some bolded text looked like hypertext but was not responsive when users clicked on it.	Mechanisms	4	4 (33)
24	Clicking on *psychological aspects of sexual pain* takes the user to the *Symptoms* page.	Treatment	3	3 (25)
25	Clicking on *learn more about how the nervous system and low arousal contribute to painful sex* takes the user to the *Symptoms* page.	Treatment	12	11 (92)
26	Clicking on *pain types* takes the user to the *Symptoms* page.	Pain types	5	4 (33)
27	Users think *anticipation of pain cycle* is located in *pain types*.	Home	2	2 (17)
28	The information underneath each section is not apparent until the section is opened.	Home	6	4 (33)
29	Links to treatment options are currently limited to only the image and not the entire box where the image is located.	Treatment	11	9 (75)
30	The website is 1-sided in favor of female partners.	Several	5	5 (42)

Table 3. Suggested design solutions.

Number	Problem	Suggested solution
6	In-text references are not directly linked to the reference list. Users have to scroll up and down in search of references.	Clicking on a reference should take the user directly to that reference.
7	Treatment pop-ups are too small. Not convenient to users.	Pop-ups should open on a new page.
13	Too much content in treatment options and mechanisms.	Bullet points are preferred.
16	Meaning on labels not immediately apparent to users.	Automatically display label meanings when hovering around the image.
20	Content on treatment pop-ups is too cluttered.	Bullet points are preferred.
21	Patient perspectives or voices are lacking on the website. It is medically oriented.	Include psychosocial aspect of sexual pain.
28	The information underneath each section is not apparent until the section is opened.	A drop-down menu under each section is preferred.
29	Links to treatment options are currently limited to only the image and not the entire box where the image is located.	Extend the link to the entire box.
30	The website is 1-sided in favor of female partners.	Include male images.

Table 4. Post-Study System Usability Questionnaire.

Category and item	Score[a], mean (SD)
System usefulness	
Overall, I am satisfied with how easy it is to use this website.	2.44 (0.93)
It was simple to use this website.	2.00 (0.78)
I was able to complete the task and scenarios quickly using this website.	2.38 (1.02)
I felt comfortable using this website.	1.88 (0.51)
It was easy to learn to use this website.	1.88 (0.66)
I believe I can know about sexual pain quickly using this website.	2.25(1.05)
Mean overall score	2.13 (0.68)
Information quality	
The website gave me error messages that told me that something went wrong.	4.63 (1.23)
Whenever I made a mistake using the system, I could recover easily and quickly.	3.00 (1.01)
The information provided on the website is clear.	2.13 (0.98)
It is easy to find the information I need.	2.25 (0.67)
The information is effective in helping me complete the tasks and scenarios.	2.38 (0.94)
The organization of the information on the screen is clear.	2.13 (0.77)
Mean overall score	2.36 (0.78)
Interface quality	
The user interface of this website was pleasant.	1.75 (0.48)
I would like to use this website.	2.25 (0.73
This website has all the functions and capabilities I expect it to have.	2.75 (1.03)
Overall, I am satisfied with this website.	2.38 (0.78)
Mean overall score	2.28 (0.56)

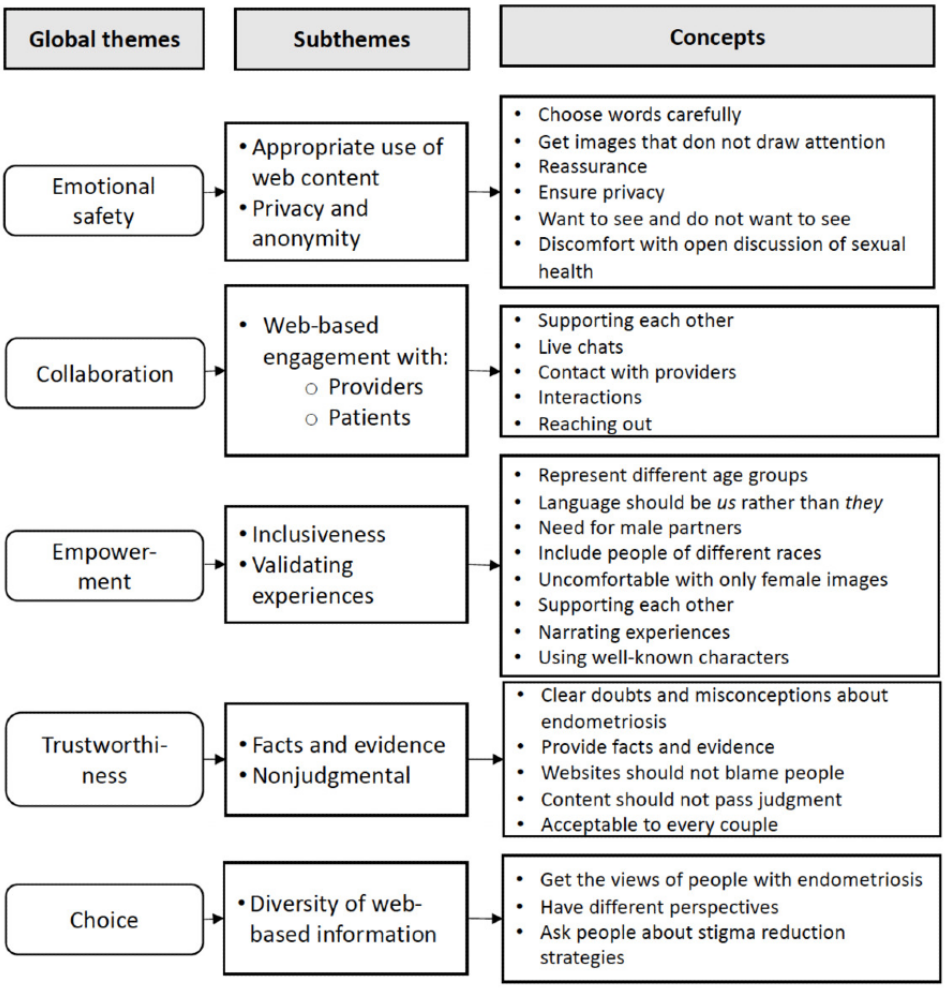
[a]Lower scores indicate better metrics of usability.

Destigmatizing Properties of Sexual Health–Related Websites

Overview

Although the focus of the stigma analysis was to identify generalities about designing destigmatizing sexual health–related websites, the participants largely referred to our website to illustrate their points. We identified participant responses that fit within the 5 main principles of the trauma-informed care framework by Fallot and Harris [25]. The 5 principles of trauma-informed care represent the global themes under which various subthemes and concepts emerged. These themes represent the participants' perspectives on how to design destigmatizing sexual health–related websites. Figure 3 shows the data analysis structure based on the trauma-informed care framework.

Figure 3. Global themes, subthemes, and concepts based on trauma-informed care.



Theme 1: Emotional Safety

Overview

According to Fallot and Harris [25], the principle of emotional safety denotes that both the setting and interaction within services are psychologically harmless, comfortable, and easy to use. In other words, emotional safety means having an awareness of individuals' discomfort or unease in using services. On the basis of this conceptualization of emotional safety, most participants in our study viewed websites as platforms that can promote emotional safety by ensuring that users are not restigmatized or retraumatized when using sexual health–related websites. In total, 2 subthemes emerged describing how stigma may be addressed through emotional safety.

Appropriate Use of Web Content

An area of concern among the participants related to how images and content are displayed on web platforms. Many of the participants indicated that the display of sexual health–related content, particularly images, determines whether they will use a web platform. Others argued that the use of websites with explicit images will likely draw people's attention to what the participants are looking at, a phenomenon that many participants disliked. Participant 3, a woman aged 45 years, stated the following:

> *When I heard this was a sexual pain website, I thought I was going to see some horrible stuff...you know what I mean...but that's not the case. You have chosen your words carefully and I wouldn't shy away from browsing this website, even on a bus.*

Owing to the potential for some web images to foment stigma, a participant suggested using images or words that would not easily draw people's attention, especially in public places.

Privacy and Anonymity

Another strategy the participants identified as a means to ensure emotional safety via sexual health–related websites was to safeguard the privacy and anonymity of people who use web-based platforms. Several participants desired privacy and anonymity, indicating that they might feel uncomfortable disclosing their information to health care providers or peers. Participant 1, a woman aged 37 years, reported the following:

> *I would not want to discuss with my doctor or family members because of the stigma. I mean...you know how people behave when it comes to sexual health issues...So I don't think I will feel comfortable discussing some of the issues you have on this website with my care provider, but I think this website can provide me with what I want without exposing myself.*

The participants generally agreed that educational websites such as ours could enable people to access needed information without necessarily disclosing their private information or their identity. Participant 10, a woman aged 51 years, specifically described the following:

> *Websites should be able to protect peoples' identity by using anonymous identities during an online interaction with others or health providers.*

Theme 2: Collaboration

Overview

Collaboration denotes partnerships and the recognition that healing happens in relationships and in a meaningful sharing of power, problems, and experiences [25]. Using our website as a reference, several participants indicated that web platforms can provide a novel opportunity for collaboration among people without revealing their identity. This form of collaboration can be either with people with similar conditions or with health care professionals. Collaboration was regarded as important because many participants indicated that people with sexual pain tend to experience pain alone, in isolation, and with limited or no opportunities to discuss or share their experiences with others. A main subtheme emerged under collaboration.

Web-Based Engagement

The desire to interact with peers via websites was a dominant theme described by most participants. For instance, participants 2 and 8 suggested the use of web-based chat rooms to promote active engagement among people with similar problems or to engage with health care providers for additional support. In reference to our website, participant 2 specifically stated the following:

> *Maybe find a way of connecting people, like, have a place where people would want to hear from each other.*

More than half of the participants suggested that contact information for health professionals could also be integrated into sexual health–related websites so that people could always establish an anonymous connection with health care providers for additional information that may not be contained on web platforms. This was specifically echoed by participant 11, a woman aged 39 years:

> *If you have the contacts of people or those behind the website where I can contact them in case I need personalized information would be useful...don't you think this will enhance the credibility and relevance of the information?*

Although several participants generally agreed on the potential benefit of some form of web-based engagement, a few others were worried about the potential privacy risk of web-based engagement. A participant contended that a person's identity could inadvertently be revealed through such engagements.

Theme 3: Empowerment

Overview

In the context of trauma-informed care, empowerment is the recognition that an individual's strengths are acknowledged, built on, and validated [25]. In other words, empowerment is aimed at conveying a sense of optimism and hope or expanding resources or an individual's capabilities. When asked how websites may help address stigma, several participants reported strategies that reflected the principle of empowerment by Fallot and Harris [25]. These participants described empowering actions that could promote a sense of optimism and hope among people who use sexual health–related websites. These empowering strategies were categorized under the 2 main subthemes of inclusiveness and validating personal experiences.

Inclusiveness

Almost all the participants noted that inclusiveness of people from diverse ethnicities, age groups, and sexes and genders on sexual health–related web platforms is an important strategy to address sexual health–related stigma. For instance, there was consensus across several participants that the presence of partners in this type of web platform was crucial in promoting emotional well-being and addressing stigma-related concerns. Approximately 6 (50%) of the participants suggested including images of partners as well as having a resource section on what partners can do to support and empower their partners who experience sexual pain. Citing our website as an example, participant 4, a woman aged 33 years, voiced the following:

> *I definitely think you need to get...partners involved in this as well because I think that's quite important...you [are] definitely not gonna fix this problem or make things easier if one party is trying to do this in isolation without some perspectives from the other party or why they can't understand what the other party is experiencing...so I think you definitely need to get that involvement in there.*

Validating Personal Experiences

Several participants noted validating each other's experiences as an empowering strategy that may help address stigma. Participant 6, for instance, expressed the desire for a section on websites where participants can narrate their experiences as a

way of empowering other people who have similar conditions. With specific reference to our website, she noted the following:

> *I think having a section where I can read about other people's stories or hear someone narrate their experiences may be helpful to me.*

Similarly, participant 4 indicated the following:

> *...having personal experiences on websites can help people who are living in isolation because such experiences can help them overcome stigmatizing problems that they otherwise would not want other people to know about.*

Validating other people's experiences was considered by many as a way to make people feel understood when using web-based platforms. A few participants also felt that having their experiences validated by others would mean they were not experiencing it in isolation.

Theme 4: Trustworthiness

Overview

Trustworthiness denotes clarity in tasks and information, accuracy, consistency, transparency, and interpersonal boundaries [25]. Applying this principle in the context of health technologies, many participants reported strategies that they thought could help dispel the myths and misconceptions regarding the causes and possible effects of endometriosis and sexual pain. For instance, participant 3 indicated that the "mere existence of a website that accurately explains endometriosis and sexual pain could help dispel the myths and misconceptions that often fuel stigma." Participant 6, a woman aged 39 years, also indicated the following:

> *A website such as this will clear a lot of doubts that many people are having regarding endometriosis and how sexual pain comes about.*

The facts and evidence on our website were also seen as ways of promoting trust in web-based content. The use of factual content and nonjudgmental information were the 2 subthemes that emerged under the global theme of trustworthiness.

Factual Content

Several participants viewed factual content as a way of addressing stigma via websites. They specifically noted that the information on our website was very factual and based on available evidence, as seen in the references. A participant suggested that other sexual health–related websites should emulate the way our website provided factual content. Participant 6 emphasized the following:

> *Just as I have said before, the information on all websites should be backed by evidence like you have here. I can see you have references to some of your claims so I can know that there is some credibility to what I am reading.*

Although several participants were positive about the facts on the website, a few others viewed the website as too medically oriented. In other words, the website did not portray patients' perspectives on endometriosis-associated sexual pain. For instance, participant 2 noted that "you only presented the facts without highlighting the psychosocial aspects or patient experiences of sexual pain."

Nonjudgmental Information

Several participants commented on and appreciated the nonjudgmental nature of the information on the sexual health–related website. With reference to our website, participant 2, a woman aged 32 years, reported the following:

> *Your message does not seem to cast doubts or pass judgment on people who suffer from endometriosis so it can be trusted by people.*

The nonjudgmental nature of the information was seen by participants as key to addressing the stigma that results from othering and stereotyping of people who live with an illness. Participant 4 indicated the following:

> *The information is not personal in nature so I don't think it may be stigmatizing. I don't think it also stereotypes anyone.*

Theme 5: Choice

Overview

Choice is the recognition of an individualized approach while strengthening people's experiences of options in services [25]. In this study, approximately 6 (50%) of the participants indicated ways in which a website might provide diverse information that maximizes an individual's choices or the diversity of options from which they can choose. A subtheme that emerged under choice related to the diversity of treatment information and opinions on web-based platforms.

Diversity of Web-Based Information

Approximately 6 (50%) of the participants indicated that, for a website to address stigma, it should facilitate people's choices by including diverse information on the treatments, supports, and resources for endometriosis used by a range of people. Awareness of diverse approaches to treating and managing endometriosis and sexual pain was considered by many to be essential to understanding the various approaches and options they might seek further information about or even consider. Considering the private nature of endometriosis and painful sex, there were currently few opportunities to learn from others about what they had tried, both successfully and unsuccessfully. Participant 7, a woman aged 42 years, specifically saw a benefit in asking people who have endometriosis about how they manage stigma related to sexual pain. She said that "asking these people about stigma will help you develop something that can address the problem." Participant 5 indicated that having different perspectives on the website will maximize people's choice of sexual pain and stigma management strategies:

> *Maybe my final suggestion will be to also get different views of people suffering from sexual pain into the website. I think it's good to hear from different perspectives of how people manage the problem.*

Discussion

Principal Findings

This study aimed to assess the usability of the *Sex, Pain, and Endometriosis* website and assess for destigmatizing properties of sexual health–related websites in general. The usability findings revealed that, except for the absence of a search bar, the possible confusion with the *mechanisms* tab, and the small pop-up windows, the participants generally found the website to be simple, uncluttered, quite easy to use, and satisfying. The system usefulness, information quality, and interface quality scores on the subscales of the PSSUQ were <3 on average, indicating good usability and satisfaction with the website. It is possible that the system usefulness, information quality, and interface quality were generally perceived as positive as the information was presented in plain language and the website was quite basic so as not to pose usability challenges. In addition to the good PSSUQ ratings, during the posttest interviews, the participants also perceived the content of the website to be credible, evidence-based, nonjudgmental, and appropriate for the age group most affected by endometriosis-associated dyspareunia. The use of nonjudgmental and age-appropriate content on the website confirms the findings of previous studies [27]. The positive findings in this study are in sharp contrast to a review that found the content of 54 endometriosis websites to be fairly inaccurate, of poor quality, noncredible, and fairly difficult to read [10]. The participants also provided suggestions and recommendations, including improving the visibility of the homepage information, explaining endometriosis and dyspareunia on the home page, and providing a drop-down menu under each section heading to make it easier to find content. Despite the positive findings, the largely minor usability problems encountered suggested the need for some revisions and redesign before the website was launched. The participants were unsatisfied with the use of text-heavy information on the website. These findings also remind us that, although more text may be needed to explain certain concepts, the frequent use of dense text on educational websites may not be favorable. Alternatively, images and bullet points could be more engaging to the participants. The recommendations from the research participants informed the revision of the website in ways that better met the needs of potential users. This evaluation study indicated the importance of user feedback in designing patient-centered educational websites on sensitive and intimate health topics [28].

To the best of our knowledge, this work is the first usability evaluation to assess for destigmatizing properties of sexual health–related web-based resources. Even though the participants found the website to be generally nonstigmatizing, the findings of the stigma analysis suggest that the website did not fully use the necessary stigma-alleviating strategies that can empower people to address stigma. For instance, the absence of male partners and inadequate information on patients' personal experiences were seen as nonempowering. The limitations of our website in addressing stigma could be attributed to the fact that we did not systematically incorporate stigma prevention into our design process even though it emerged during our needs assessment and landscape analysis. Despite these (minimal) setbacks, this study provides preliminary evidence that suggests a trauma-informed approach may inform strategies that can help web designers in developing destigmatizing websites. This is important as web platforms are increasingly used to disseminate information on sensitive health topics [11]. However, no strategies or guiding principles exist to help designers address stigma on websites despite stigma being a major issue in sexual health–related topics. As trauma and stigma are inherently intertwined and work to reinforce each other [26,29], these findings suggest that adopting a trauma-informed approach to developing digital platforms may help address stigma concerns among users of web-based resources. For instance, the anonymous web-based channels identified under the principle of collaboration may motivate people to reach out when in need of support, provide avenues for web-based consultation or engagement with health professionals, and provide a channel for sharing coping mechanisms and supporting each other to overcome the stigma of pain symptoms. The findings of this study demonstrate how other studies that have adopted a trauma-informed approach to successfully design interventions have helped address stigma concerns among people living with HIV and AIDS in other settings [29].

This study also shows how inclusiveness can be applied in web design to address sexual health–related stigma. Inclusive design approaches suggest all-encompassing ways in which websites could be optimized to be usable and acceptable to diverse populations with respect to ethnicity, gender, age, and other forms of human differences. Although inclusiveness was a concern among the patient partners during our needs assessment, we cannot say that our website fully maximized the principles of inclusiveness as the participants noted the absence of male partners during the usability analysis. Future websites on sensitive health topics should start with the principles of inclusive design from the outset. With the principle of inclusiveness, this study extends the emphasis on inclusive design principles from a predominant focus on older adults and people with disabilities to include diverse human variabilities such as ethnicity, gender, and diversity in perspectives [30].

Although the various strategies identified by the participants could help in creating destigmatizing websites, some of the strategies may have an inherent privacy risk. In other words, some of the strategies may reveal rather than conceal a user's identity to others. For instance, some participants were worried about the privacy risk of web-based communication features such as chat rooms. This worry was not surprising, as the use of chat rooms and other two-way web-based communication features has been associated with privacy breaches in other studies [31]. The participants were not particularly worried about our website as it does not offer any web-based two-way communication features. However, they expressed privacy concerns for other sexual health–related websites that collect personal information. These findings reflect studies that suggest that anonymous websites or websites that do not collect or store any personal information, including ours, may be useful for obtaining information on stigmatized conditions or symptoms such as sexual pain [32,33]. However, the participants' suggestions for anonymous chat rooms on websites may be an

avenue for abusive and web-based stigmatization by peers [16]. This can negate the importance of this potentially safe space. The potential privacy and security risks inherent in chat rooms suggest the need for special design considerations to make anonymous chat rooms or live chats safer.

Given that people may be exposed to all kinds of information from varying sources, the facts and evidence-based information on our website were seen as credible and reliable information that will ultimately address the misconceptions of endometriosis and sexual pain. Although there is still room for improvement, our website also ensures a full range of choices by showing different treatment options for endometriosis-associated dyspareunia. What was missing to fully maximize the principle of choice was having diverse patient perspectives captured on the platform. Our team grappled with providing information about self-management or treatment strategies for which there is only anecdotal evidence, deciding in the end to highlight the information that has greater theoretical or empirical evidence to support its use. Furthermore, the desires for diverse information and for factual information (under the principle of trustworthiness) were very interesting but contradictory to each other, though consistent with our team's internal debate while creating the website. Although the diversity of web-based information was considered essential to understanding the various options and approaches that people with endometriosis-associated sexual pain might take, evidence may not be available to support all the different options people may desire.

The need for websites to convey real-life experiences and facilitate interaction among people as a way of addressing sexual health–related stigma is consistent with previous studies [17,34]. These previous studies demonstrated how the use of positive language and patients' personal experiences can promote the uptake of web-based resources. These findings also show that our website may be particularly useful for people who need evidence-based information on endometriosis-associated dyspareunia but are fearful of disclosing or discussing their symptoms with providers.

Limitations

Instead of in-person usability testing, we opted for remote usability testing to comply with public health orders during the COVID-19 pandemic. We turned off the video to protect the users' privacy and, therefore, could not observe body gestures and facial expressions that typically convey participants' reactions during a think-aloud procedure [35]. However, given the sensitive nature of the topic, remote usability testing might have turned out to be a blessing in disguise as the participants had the opportunity to explore the website from their home and in relative anonymity. Although the participants emphasized the important role of partners in managing sexual health–related stigma, our sample included only people diagnosed with endometriosis. This limitation is consistent with the broad sexual health literature, where partner engagement in women's sexual health research is particularly challenging and, in some cases, disregarded [36-38]. Future usability studies on websites that address sexual health problems should include partners to ensure an inclusive end product. For an educational website on endometriosis-associated dyspareunia such as this one, partner perspectives are particularly important as the psychosocial impact of sexual pain can also be experienced by partners [39].

Conclusions

Websites on sensitive health topics are increasingly designed to provide anonymous platforms for people to obtain evidence-based information and to empower users, dispel myths, and alleviate stigma.

Acknowledgments

Funding for this project was received from the Canadian Institutes of Health Research Project Grant Competition (grant PJT-156084). The funding sources had no involvement in the study design, data collection, analysis, writing of the report, or the decision to publish this manuscript.

Conflicts of Interest

None declared.

References

1. Leyland N, Casper R, Laberge P, Singh SS, Allen L, Arendas K, et al. Endometriosis: diagnosis and management. J Endometr 2018;2(3):107-134 [FREE Full text] [doi: 10.1177/228402651000200303]
2. Giudice LC. Clinical practice. Endometriosis. N Engl J Med 2010;362(25):2389-2398 [FREE Full text] [doi: 10.1056/NEJMcp1000274] [Medline: 20573927]
3. Caruso S, Iraci M, Cianci S, Casella E, Fava V, Cianci A. Quality of life and sexual function of women affected by endometriosis-associated pelvic pain when treated with dienogest. J Endocrinol Invest 2015;38(11):1211-1218. [doi: 10.1007/s40618-015-0383-7] [Medline: 26337183]
4. Wahl KJ, Orr NL, Lisonek M, Noga H, Bedaiwy MA, Williams C, et al. Deep dyspareunia, superficial dyspareunia, and infertility concerns among women with endometriosis: a cross-sectional study. Sex Med 2020;8(2):274-281 [FREE Full text] [doi: 10.1016/j.esxm.2020.01.002] [Medline: 32061579]
5. Yong PJ. Deep dyspareunia in endometriosis: a proposed framework based on pain mechanisms and genito-pelvic pain penetration disorder. Sex Med Rev 2017;5(4):495-507. [doi: 10.1016/j.sxmr.2017.06.005] [Medline: 28778699]

6. Ferrero S, Esposito F, Abbamonte LH, Anserini P, Remorgida V, Ragni N. Quality of sex life in women with endometriosis and deep dyspareunia. Fertil Steril 2005;83(3):573-579. [doi: 10.1016/j.fertnstert.2004.07.973] [Medline: 15749483]
7. Di Donato N, Montanari G, Benfenati A, Monti G, Bertoldo V, Mauloni M, et al. Do women with endometriosis have to worry about sex? Eur J Obstet Gynecol Reprod Biol 2014;179:69-74. [doi: 10.1016/j.ejogrb.2014.05.022] [Medline: 24965983]
8. Denny E. Women's experience of endometriosis. J Adv Nurs 2004;46(6):641-648. [doi: 10.1111/j.1365-2648.2004.03055.x] [Medline: 15154905]
9. Weijts W, Houtkoop H, Mullen P. Talking delicacy: speaking about sexuality during gynaecological consultations. Sociol Health Illn 1993;15(3):295-314. [doi: 10.1111/1467-9566.ep10490531]
10. Hirsch M, Aggarwal S, Barker C, Davis CJ, Duffy JM. Googling endometriosis: a systematic review of information available on the Internet. Am J Obstet Gynecol 2017;216(5):451-8.e1. [doi: 10.1016/j.ajog.2016.11.1007] [Medline: 27840143]
11. Redston S, de Botte S, Smith C. Resolving embarrassing medical conditions with online health information. Int J Med Inform 2018;114:101-105. [doi: 10.1016/j.ijmedinf.2018.03.010] [Medline: 29605386]
12. Berger M, Wagner TH, Baker LC. Internet use and stigmatized illness. Soc Sci Med 2005;61(8):1821-1827. [doi: 10.1016/j.socscimed.2005.03.025] [Medline: 16029778]
13. Strategy for patient-oriented research. Canadian Institutes of Health Research. 2019. URL: http://www.cihr-irsc.gc.ca/e/45851.html [accessed 2020-08-12]
14. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
15. Ho K, Jarvis-Selinger S, Lauscher HN, Cordeiro J, Scott R. Technology enabled knowledge translation for eHealth: principles and practice. New York: Springer; 2012.
16. Bauermeister JA, Muessig KE, LeGrand S, Flores DD, Choi SK, Dong W, et al. HIV and sexuality stigma reduction through engagement in online forums: results from the HealthMPowerment intervention. AIDS Behav 2018;23(3):742-752. [doi: 10.1007/s10461-018-2256-5] [Medline: 30121727]
17. Magee JC, Bigelow L, Dehaan S, Mustanski BS. Sexual health information seeking online: a mixed-methods study among lesbian, gay, bisexual, and transgender young people. Health Educ Behav 2012;39(3):276-289. [doi: 10.1177/1090198111401384] [Medline: 21490310]
18. Klein C, Lomonaco C. Real talk: developing a computer-delivered sexual health program for black men who have sex with men. AIDS Educ Prev 2016;28(6):455-471. [doi: 10.1521/aeap.2016.28.6.455] [Medline: 27925486]
19. Davis WM, Shoveller JA, Oliffe JL, Gilbert M. Young people's perspectives on the use of reverse discourse in web-based sexual-health interventions. Cult Health Sex 2012;14(9):1065-1079. [doi: 10.1080/13691058.2012.714800] [Medline: 22943601]
20. Abdulai AF, Howard AF, Currie LM. Stigmatizing and de-stigmatizing properties of Web apps for sexual health-related conditions: a scoping review. Int J Sex Health 2021:1-16. [doi: 10.1080/19317611.2021.1949655]
21. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J Hum-Comput Interact 1995;7(1):57-78. [doi: 10.1080/10447319509526110]
22. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, et al. Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2010;12(3):e30 [FREE Full text] [doi: 10.2196/jmir.1349] [Medline: 20675293]
23. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Harris M, Fallot RD. Envisioning a trauma-informed service system: a vital paradigm shift. New Dir Ment Health Serv 2001(89):3-22. [doi: 10.1002/yd.23320018903] [Medline: 11291260]
26. Frieh EC. Stigma, trauma and sexuality: the experiences of women hospitalised with serious mental illness. Sociol Health Illn 2020;42(3):526-543. [doi: 10.1111/1467-9566.13034] [Medline: 31758579]
27. Parasar P, Ozcan P, Terry KL. Endometriosis: epidemiology, diagnosis and clinical management. Curr Obstet Gynecol Rep 2017;6(1):34-41 [FREE Full text] [doi: 10.1007/s13669-017-0187-1] [Medline: 29276652]
28. Orenstein EW, Boudreaux J, Rollins M, Jones J, Bryant C, Karavite D, et al. Formative usability testing reduces severe blood product ordering errors. Appl Clin Inform 2019;10(5):981-990 [FREE Full text] [doi: 10.1055/s-0039-3402714] [Medline: 31875648]
29. Whetten K, Reif S, Toth M, Jain E, Leserman J, Pence BW. Relationship between trauma and high-risk behavior among HIV-positive men who do not have sex with men (MDSM). AIDS Care 2012;24(11):1453-1460 [FREE Full text] [doi: 10.1080/09540121.2012.712665] [Medline: 22909318]
30. Newell AF, Gregor P, Morgan M, Pullin G, Macaulay C. User-sensitive inclusive design. Univ Access Inf Soc 2010;10(3):235-243. [doi: 10.1007/s10209-010-0203-y]
31. Ralph LJ, Berglas F, Schwartz SL, Brindis CD. Finding teens in TheirSpace: using social networking sites to connect youth to sexual health services. Sex Res Soc Policy 2011;8(1):38-49. [doi: 10.1007/s13178-011-0043-4]

32. Gannon B, Davis R, Kuhns LM, Rodriguez RG, Garofalo R, Schnall R. A mobile sexual health app on empowerment, education, and prevention for young adult men (MyPEEPS mobile): acceptability and usability evaluation. JMIR Form Res 2020;4(4):e17901 [FREE Full text] [doi: 10.2196/17901] [Medline: 32254043]
33. Muessig KE, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman LB. A systematic review of recent smartphone, Internet and Web 2.0 interventions to address the HIV continuum of care. Curr HIV/AIDS Rep 2015;12(1):173-190. [doi: 10.1007/s11904-014-0239-3] [Medline: 25626718]
34. Lagu T, Norton CM, Russo LM, Priya A, Goff SL, Lindenauer PK. Reporting of patient experience data on health systems' websites and commercial physician-rating websites: mixed-methods analysis. J Med Internet Res 2019;21(3):e12007 [FREE Full text] [doi: 10.2196/12007] [Medline: 30916654]
35. Alhadreti O, Mayhew P. Rethinking thinking aloud: a comparison of three think-aloud protocols. In: Proceedings of the 2018 CHI conference on human factors in computing systems. 2018 Presented at: CHI'18; April 21-26, 2018; Montreal p. 1-12. [doi: 10.1145/3173574.3173618]
36. Davis AC, Wright CJ, Temple-Smith MJ, Hellard ME, Lim MS. A health education website developed to meet young people's information needs about Web-based pornography and sharing of sexually explicit imagery (SCOPE): usability study. JMIR Form Res 2019;3(3):e12824. [doi: 10.2196/12824] [Medline: 31411140]
37. Richters J, Badcock PB, Simpson JM, Shellard D, Rissel C, de Visser RO, et al. Design and methods of the Second Australian Study of Health and Relationships. Sex Health 2014;11(5):383-396. [doi: 10.1071/SH14115] [Medline: 25376992]
38. Davis AC, Carrotte ER, Hellard ME, Lim MS. What behaviors do young heterosexual Australians see in pornography? A cross-sectional study. J Sex Res 2018;55(3):310-319. [doi: 10.1080/00224499.2017.1417350] [Medline: 29333877]
39. Facchin F, Barbara G, Saita E, Mosconi P, Roberto A, Fedele L, et al. Impact of endometriosis on quality of life and mental health: pelvic pain makes the difference. J Psychosom Obstet Gynaecol 2015;36(4):135-141. [doi: 10.3109/0167482X.2015.1074173] [Medline: 26328618]

Abbreviations

PSSUQ: Post-Study System Usability Questionnaire

Edited by A Kushniruk; submitted 18.06.21; peer-reviewed by C Klein, D Karavite, M Salimi; comments to author 26.08.21; revised version received 18.10.21; accepted 06.11.21; published 03.03.22.

Please cite as:
Abdulai AF, Howard AF, Yong PJ, Noga H, Parmar G, Currie LM
Developing an Educational Website for Women With Endometriosis-Associated Dyspareunia: Usability and Stigma Analysis
JMIR Hum Factors 2022;9(1):e31317
URL: https://humanfactors.jmir.org/2022/1/e31317
doi:10.2196/31317
PMID:35238789

Original Paper

Using Health Concept Surveying to Elicit Usable Evidence: Case Studies of a Novel Evaluation Methodology

Alex Mariakakis[1], PhD; Ravi Karkar[2], PhD; Shwetak N Patel[2], PhD; Julie A Kientz[3], PhD; James Fogarty[2], PhD; Sean A Munson[3], PhD

[1]Department of Computer Science, University of Toronto, Toronto, ON, Canada

[2]School of Computer Science & Engineering, University of Washington, Seattle, WA, United States

[3]Department of Human Centered Design & Engineering, University of Washington, Seattle, WA, United States

Corresponding Author:
Alex Mariakakis, PhD
Department of Computer Science
University of Toronto
Bahen Centre for Information Technology, Room 4283
40 St. George Street
Toronto, ON, M5S 2E4
Canada
Phone: 1 416 978 5359
Email: mariakakis@cs.toronto.edu

Abstract

Background: Developers, designers, and researchers use rapid prototyping methods to project the adoption and acceptability of their health intervention technology (HIT) before the technology becomes mature enough to be deployed. Although these methods are useful for gathering feedback that advances the development of HITs, they rarely provide usable evidence that can contribute to our broader understanding of HITs.

Objective: In this research, we aim to develop and demonstrate a variation of vignette testing that supports developers and designers in evaluating early-stage HIT designs while generating usable evidence for the broader research community.

Methods: We proposed a method called *health concept surveying* for untangling the causal relationships that people develop around conceptual HITs. In health concept surveying, investigators gather reactions to design concepts through a scenario-based survey instrument. As the investigator manipulates characteristics related to their HIT, the survey instrument also measures proximal cognitive factors according to a health behavior change model to project how HIT design decisions may affect the adoption and acceptability of an HIT. Responses to the survey instrument were analyzed using path analysis to untangle the causal effects of these factors on the outcome variables.

Results: We demonstrated health concept surveying in 3 case studies of sensor-based health-screening apps. Our first study (N=54) showed that a wait time incentive could influence more people to go see a dermatologist after a positive test for skin cancer. Our second study (N=54), evaluating a similar application design, showed that although visual explanations of algorithmic decisions could increase participant trust in negative test results, the trust would not have been enough to affect people's decision-making. Our third study (N=263) showed that people might prioritize test specificity or sensitivity depending on the nature of the medical condition.

Conclusions: Beyond the findings from our 3 case studies, our research uses the framing of the Health Belief Model to elicit and understand the intrinsic and extrinsic factors that may affect the adoption and acceptability of an HIT without having to build a working prototype. We have made our survey instrument publicly available so that others can leverage it for their own investigations.

(JMIR Hum Factors 2022;9(1):e30474) doi:10.2196/30474

KEYWORDS

mobile health; survey instrument; health screening; health belief model; path analysis; user design; health technology; health intervention technology; digital health; mobile phone

Introduction

Overview

There are numerous design decisions beyond the rigor of the information being presented in a health intervention technology (HIT) that can affect how people incorporate the HIT's guidance into their decision-making [1]. These factors can range from the HIT's visual appearance [2] and message framing [3,4] to people's beliefs and psychological traits [5,6]. Late-stage evaluation methods such as A/B field testing and randomized controlled trials are designed to help HIT creators explore the ways in which the aforementioned factors might affect people's decision-making [7-12]. However, deploying an HIT too early can expose people to numerous risks, such as delays in necessary lifestyle changes, postponed diagnoses, and unwarranted stress. User-centered design also encourages designers to incorporate feedback early and often in their process before reaching these late-stage evaluation methods [13]. Unfortunately, early-stage evaluation and rapid prototyping methods (eg, think-aloud evaluations and paper prototyping) are not as well-suited for eliciting feedback on how people would respond to an HIT's guidance. Many people assess the credibility of an HIT based on its visual appearance and language [2,14], which may not be fully developed in a low-fidelity prototype. People can also idealize unspecified HIT features to their liking, resulting in a positive but biased evaluation [15]. Even when a prototype is complete, early-stage methods are better suited for identifying *which* features people prefer but not *why* they prefer those features or *how* those features will affect use [16].

In light of these challenges, Klasnja et al [17] called for early-stage evaluation methods that generate *usable evidence*: "empirical findings about the causal effects of [HITs] and how those effects vary with individual differences, context of use, and system design." Klasnja et al [17] discussed usable evidence in the context of developers and designers who are creating a new HIT; however, there is also a broader need within the research community to generate findings that lead to guidelines and theories. Identifying usable evidence requires an explicit understanding of the causal mechanisms that affect the reception of an HIT [18], which can only be gained by untangling the effects of HIT design decisions and proximal cognitive factors such as beliefs and attitudes.

As a methodological contribution to HIT design research, we propose *health concept surveying*, a variation of vignette testing [19,20] that supports the generation of usable evidence. Health concept surveying is centered on a survey instrument that presents target users with a technology concept in a scenario and then measures the potential impact that HIT design decisions may have on 2 distal outcomes [21,22]: (1) adoption of an HIT, which is a person's intention of using an HIT, and (2) acceptability of an HIT's suggestions, which is a person's willingness to conduct the follow-up actions recommended by the HIT.

The survey instrument also measures proximal cognitive factors as defined by a health behavior change framework (eg, the Health Belief Model, HBM [23,24]). The responses to the survey were analyzed using path analysis to surface causal pathways that inform future research on HITs. As health concept surveying relies on design concepts rather than physical prototypes, HIT creators can be selective about which HIT design characteristics they include to prevent study participants from getting distracted by missing or incomplete features.

We demonstrate the efficacy of health concept surveying using 3 case studies to display its utility for multiple stakeholders. The first 2 case studies show how health concept surveying would be beneficial to a developer or designer invested in a particular HIT, whereas the third case study highlights how researchers could use health concept surveying to test a broader hypothesis across multiple HITs. The case studies are centered on sensor-based health-screening apps—smartphone apps that use on-device sensors such as cameras and microphones to identify the presence of medical symptoms—as this domain is emerging in academia and industry alike [25]. The design decisions that are explored in these case studies include (1) the inclusion of an incentive, (2) the inclusion of visual test result explanations, and (3) the trade-off between the true positive rate and true negative rate.

In summary, our research contributes the following:

1. The health concept surveying method, which uses vignette testing to disentangle the effects of design decisions and proximal cognitive factors on the adoption and acceptability of an HIT.
2. Case studies that show how health concept surveying can be used to benefit specific HIT designs while generating usable evidence for the broader community.
3. A more complex case study that shows how health concept surveying can also support more abstract research to directly contribute to our understanding of HITs.

Prior Work

Our research is primarily inspired by a collection of commentaries on behavior change technologies (BCTs) by Klasnja et al and Hekler et al [17,26,27]. BCTs aim to persuade a person to change their habits, whereas HITs can include both health-focused BCTs and technologies that provide a 1-time suggestion for a course of action.

In this thread of research, Klasnja et al [26] first recognized that demonstrating behavior change for early-stage BCTs is often "infeasible as well as unnecessary for a meaningful contribution to HCI research" and instead suggest that researchers strive for "a deep understanding of the how and why of the system use by its target users." They proposed that researchers can work toward such an understanding by tailoring their evaluation methods to the intervention strategies involved in their HIT (eg, self-monitoring, conditioning, and tunneling [28]), which can require the development of new strategies that balance abstraction with contextual relevance [27]. By leveraging behavioral science theories, Klasnja et al [17] suggested that researchers can not only advance their particular intervention but also generate *usable evidence*: "empirical findings about the causal effects of BCTs and how those effects vary with individual differences, context of use, and system design."

Evaluation methods such as factorial designs [7,8], microrandomized trials [9,10], and single-case experimental

designs [11,12] can be used to methodically test hypothesis-driven research; however, these methods are typically considered only after a prototype is sophisticated enough to be put into people's hands. By using a survey method, health concept surveying allows investigators to include as few or as many details about an HIT as they deem fit. This flexibility of abstraction not only makes the health concept surveying suitable for developers and designers with early-stage HITs but also for researchers as they explore hypotheses around HIT concepts. Health concept surveying also relies on health behavior change frameworks so that researchers can disentangle complicated relationships between factors to generate usable evidence.

Theory: HBM

Social psychologists have proposed various frameworks to predict, explain, and change health behaviors in matters related to public and personalized health. These frameworks have been applied to topics ranging from smoking cessation and exercise [29] to vaccination [30] and hearing loss prevention [31]. Health behavior change frameworks typically fall into two categories [32]: social cognition models (eg, theory of planned behavior [33] and HBM [23,24]), which use cognitive factors such as beliefs and attitudes as proximal determinants of behavior; and stage models (eg, transtheoretical model [34]), which describe decisions as a sequence of discrete phases.

Survey instruments for applying health concept surveying could be modeled after any of the aforementioned health behavior change frameworks to specify proximal cognitive factors. In this work, we demonstrate health concept surveying with a survey instrument based on the HBM. Researchers have criticized aspects of the HBM, such as its lack of applicability outside of health-related contexts [35,36] and the inconsistency in how different researchers define its constructs [35,37,38]. Nevertheless, we use the HBM because of its specific focus on health interventions, its applicability to both short-term actions and long-term behaviors, and the potential for its constructs to map to actionable feedback for developers, designers, and researchers. By providing a survey instrument that others can use, we hope to provide standardized questions that mitigate inconsistency.

The HBM posits that a person will undergo an action to improve or maintain their health if the perceived barriers to that particular action are outweighed by the perceived seriousness of the health problem, the perceived susceptibility to that health problem, and the perceived benefits of taking action. All of these constructs are affected by modifying variables, that is, demographic information and psychological characteristics that can explain a person's decision-making. For instance, someone who is well-educated may understand the benefits of early screening, whereas someone who does not have flexible income may view the cost of a screening examination as burdensome. Conceptually, the HBM can be summarized using the following equation:

$$\textit{Modifying variables} \times (\textit{Seriousness} + \textit{Susceptibility} + \textit{Benefits} - \textit{Barriers}) + \textit{Cues to action} = \textit{Likelihood of action}$$

Definitions of the HBM constructs according to Urich [39] are provided in Textbox 1.

Textbox 1. The constructs of the Health Belief Model and their definitions.

Health Belief Model constructs and definitions

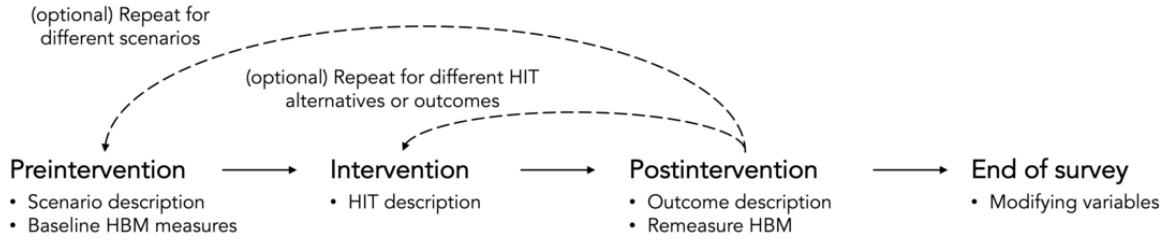
- Perceived seriousness: a person's subjective assessment of the severity of the health problem and its potential consequences
- Perceived susceptibility: a person's subjective assessment of their risk of developing the health problem
- Perceived benefits: a person's subjective assessment of the value in taking a certain action
- Perceived barriers: a person's subjective assessment of the obstacles to taking a certain action
- Modifying variables: individual characteristics (demographic and psychosocial) that can affect a person's perception of a health problem
- Cues to action: internal or external triggers that prompt a certain action

Methods

Overview

Health concept surveying is centered on a survey instrument that allows investigators to measure proximal cognitive factors while manipulating HIT features. In this section, we provide details on the contents of the survey instrument, as illustrated in Figure 1. We illustrated this survey instrument with a concept for a sensor-based health-screening app called SkinCheck, which analyzes the appearance of a person's mole to determine whether it is cancerous. A complete example of the survey instrument used can be found in Multimedia Appendix 1.

Figure 1. The structure of the survey instrument for health concept surveying comprises four stages: (1) preintervention, (2) intervention, (3) postintervention, and (4) end of survey. HBM: Health Belief Model; HIT: health intervention technology.



Survey Design

Preintervention

Our survey instrument starts by presenting respondents with a scenario that describes a cue to action related to the health topic of interest for the HIT. Cues to action can include the emergence of symptoms, promotional advertising, or even direct recommendations or prescriptions from a physician. For our example regarding a sinus infection, our prompt was as follows:

> *You recently noticed a new mole (beauty mark) on your arm that is oddly colored and misshapen. After looking up information online, you worry that you might be developing skin cancer.*

After reading the scenario, the respondent is asked to complete an instructional manipulation check (IMC) [40], where they are asked to select the symptoms that are associated with the described medical condition. In addition to checking that the respondent read the scenario, the IMC forces the respondent to spend extra time reflecting on the scenario.

The respondent is then asked a series of questions related to their initial reactions to the scenario according to the constructs of the HBM: *PerceivedSeriousness*, *PerceivedSusceptibility*, *PerceivedBenefits*, and *PerceivedBarriers* (Textbox 2). Each construct has a corresponding question except for *PerceivedSeriousness*, which has 3 questions to account for the various impacts that a health-related issue can have on a person's life. All responses are recorded on a 7-point scale. The respondent is also asked whether they would take various actions as a series of yes-or-no questions. The respondent is free to take 0, 1, or multiple actions; therefore, we use the variable *ActionType* to keep track of which action corresponds to each response and *ActionTaken* to track whether the respondent would take each action. As people can foresee different *PerceivedBenefits* and *PerceivedBarriers* for various actions, we also ask the respondent to separately rate those questions for each *ActionType*.

Textbox 2. The set of questions that are asked in the pre- and postintervention stages of the health concept surveying survey instrument.

Health Belief Model constructs and survey questions

- Perceived seriousness
  - If you had [medical condition] in this scenario, how impactful do you believe it would be on your long-term health?
  - If you had [medical condition] in this scenario, how impactful do you believe it would be on your finances?
  - If you had [medical condition] in this scenario, how impactful do you believe it would be socially and/or professionally?
- Perceived susceptibility
  - How likely do you think you are to have [medical condition] in this scenario?
- Perceived benefits
  - How beneficial do you believe each of these actions would be towards helping you recover from your symptoms?
- Perceived barriers
  - How easy do you think it would be for you to take each of the following actions to help you recover from your symptoms?
- Action taken
  - Given the possibility that you may [have/not have] [medical condition], which of the following actions would you plan to take on the same day as when you discovered your symptoms?

Intervention

After the respondents report which actions they would take, they are given information about an HIT that is meant to address the health-related issue described in the scenario. This is where the investigator can choose which details to include about their HIT. Although more details will generally make the HIT concept more concrete and leave less room for uncertainty, the investigator may choose to leave out some information to avoid potential distractions from their primary questions. Our SinusCheck example includes the following text:

> *A smartphone app named SkinCheck analyzes a picture of a mole to determine whether or not it is cancerous. To use the app, you are asked to take a picture of the mole so that it is clearly visible. The app guides you through taking a picture so that it can see the mole clearly and at a proper distance.*
>
> *SkinCheck comes with your smartphone by default as part of a new mobile health initiative by [Phone Company]. SkinCheck provides text-based and audio-based instructions to help you perform the test. The app also checks that the test was performed correctly. You can repeat the test until the app determines the image to be "valid." The results of the test are available instantly.*

This example includes a high-level description of the app's source and functionality; however, it does not include any mockups or screenshots of the app itself. Therefore, an investigator could use this example early in their development process to explore how people would feel about the concept of using an app to detect sinus infections without undue influence from the visuals of the app itself, which could be addressed at a later time.

Postintervention Stage

After reading the HIT description, the respondent is asked about their interest in using the HIT on a 7-point scale, which we call *TechnologyInterest*. If the respondent says that they would use the HIT beyond the neutral score, they are taken to pages where they are asked to react to different outcomes in a randomized order. For health-screening apps, our outcomes included positive and negative test results. After each outcome, the respondent is asked to re-evaluate their responses to the questions in Textbox 2. We can determine whether the HIT would have changed the respondent's plan by comparing *ActionTaken* across the pre- and postintervention stages. This produces a second outcome variable called *ActionChange*, indicating whether the HIT had sufficient influence to change a person's behavioral intent. Similar to *ActionTaken*, *ActionChange* is recorded for each *ActionType*.

Every HBM construct would ideally be evaluated before and after the intervention to examine how perceptions changed as a result of the intervention. However, doing so can significantly increase the survey length when evaluating multiple versions of an HIT. Therefore, an investigator may choose to remove a postintervention question for a particular HBM if they are confident that their design question is unrelated to it. In such cases, the response from the preintervention stage is propagated through the rest of the respondent's data, as it is assumed to be constant. We use this modification in our third case study as it has 3 manipulated factors and a mixed factorial study design.

End of Survey

At the end of the survey instrument, the respondent is asked for information related to *ModifyingVariables* within the HBM. These questions can capture demographic information (eg, age and access to health care services), psychological properties (eg, risk aversion), or self-assessed expertise in topics related to the HIT (eg, numeracy and familiarity with the medical condition). As the content of the survey itself can provide new information to respondents, some of these questions may be best placed at the beginning of the survey.

Design Summary

To summarize, our survey instrument captures two key outcome variables: (1) *TechnologyInterest*, which measures the likelihood that the respondent would use the app on a 7-point scale, and (2) *ActionTaken*, which measures the likelihood that the respondent would take action based on the information available to them at that point in the survey. All respondents would answer questions related to each HBM construct, *TechnologyInterest* and *ActionTaken* in the preintervention stage. Respondents who express sufficient interest in using the HIT are then shown various potential outcomes of the HIT and asked to reanswer the HBM construct and *ActionTaken* questions for each one. The responses to *ActionTaken* in the pre- and postintervention stages are compared for each HIT outcome to form the outcome variable *ActionChange*. *ActionChange* is not recorded for respondents who do not express interest in using the HIT as they never reach the postintervention stage. We use *TechnologyInterest* to project the potential adoption of an HIT, and we use *ActionChange* to project the potential acceptability of an HIT.

Analysis

We analyzed data from our survey instrument using path analysis, a variant of structural equation modeling that discerns the effects of a set of observable variables on a specified outcome via multiple causal pathways [41]. Path analysis revolves around graphical models called path diagrams, which encode hypothesized causal relationships by using nodes to represent measured constructs and directed edges to represent the relationships between them. Running path analysis produces a model in which each edge is assigned a path coefficient and a corresponding *P* value. We reported standardized path coefficient (*b*), where $b=0.5$ from *X* to *Y*, suggesting that a 1 SD change in *X* produces a 0.5 SD change in *Y*.

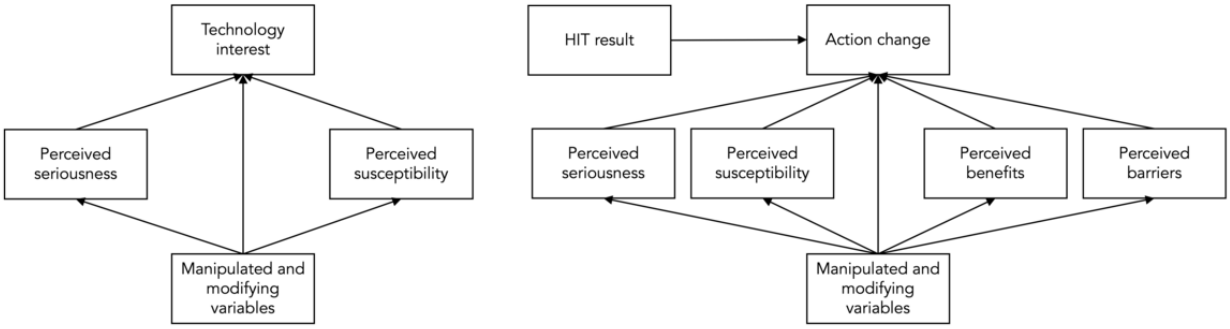
The result of path analysis is a model in which each edge in the path diagram is assigned a path coefficient and *P* value. The coefficient is not a correlation coefficient but rather indicates the degree to which one variable influences the other. Chin [42] asserted that meaningful path coefficients have absolute magnitudes >0.2. The models themselves can be assessed according to a variety of fit statistics with no agreed-upon standard [43-45]. We reported two fit statistics: comparative fit index (CFI) and standardized root mean square residual (SRMR). CFI compares the model fit against the fit of an

independent model in which the variables are assumed to be uncorrelated, whereas SRMR compares the difference between the residuals of the covariance matrix and the hypothesized covariance model while standardizing for elements with different ranges. Hu and Bentler [46] considered a model fit to be strong when its CFI is ≥0.95 and its SRMR is ≤0.09. The fit statistics are likely to be poor if the path diagram is insufficient for characterizing the relationship between variables (eg, missing nodes or edges) or if the responses to key variables are heavily biased.

It is possible to analyze the data that are gathered with our survey instrument using techniques such as analysis of variance or generalized linear models; however, separate regressions would be needed for each variable with an inbound edge to capture all the causal pathways in the path diagram. Path analysis makes it easier for investigators to contrast the importance of 2 causal relationships as the entire path diagram is processed at once, and the edge weights are directly comparable. Path analysis also makes it possible to characterize the mediated relationships. In other words, the influence of *X* on *Z* via *Y* can be calculated by multiplying the edge weights from *X* to *Y* and from *Y* to *Z*.

Figure 2 shows the diagram of our outcome variables. *PerceivedSeriousness* is a latent variable that combines the responses to its 3 constituent questions. The more nodes that are in the path diagram, the more complicated the model becomes and the more participants that must be recruited to achieve statistical significance. Therefore, we encourage investigators to remove directed edges between 2 variables if they are confident that the variables are unrelated according to their definition or the investigators' best judgment. For example, we assume that *TechnologyInterest* is independent of *PerceivedBenefits* and *PerceivedBarriers* as those constructs relate to actions that are unrelated to using the HIT itself. HIT design variables and *ModifyingVariables* should also be added at the investigators' discretion, with particular focus paid to when they are introduced in the survey instrument. If a design decision affects how the HIT is introduced, the corresponding variable should be added to both path diagrams; however, if the design decision only appears during the intervention stage, the variable should not be included in the *TechnologyInterest* diagram.

Figure 2. The basic path diagrams used to disentangle the effects that health intervention technology design decisions and user-intrinsic factors have on the measured outcome variables: TechnologyInterest (left) and ActionChange (right). HIT: health intervention technology.



We fit the *TechnologyInterest* model to the data from all respondents using their ratings for the HBM constructs in the preintervention stage. Models for *ActionChange* require using data from both the pre- and postintervention stages, therefore limiting the analysis to data from respondents who expressed sufficient interest in using the HIT. Variables such as *ActionType* do not have causal effects but still produce unique entries in the data set. Rather than including these variables in the path diagrams, they are used as grouping factors for multigroup path analysis, a technique in which a model is fit for each group with assumptions about which attributes the models share. As people can have asymmetric reactions to positive and negative test results, we fit separate *ActionChange* models in response to positive (*ActionChangePositive*) or negative (*ActionChangeNegative*) test results when applicable. In each of these cases, we excluded respondents who would have taken the HIT's *target* action in the preintervention stage. For example, respondents who would have taken action in the preintervention stage were excluded from the model because a positive test result would not be needed to convince them to take action.

Results

Overview

To demonstrate the flexibility of our method in a series of case studies, we first had to create a variety of prompts for plausible health-related scenarios and sensor-based health-screening apps. We selected three scenarios based on their plausibility and the different reactions we expected them to elicit: (1) a scenario involving pink eye, which represents a *common* medical condition; (2) a scenario involving skin cancer, which represents a *serious* medical condition; and (3) a scenario involving halitosis, which represents a *stigmatizing* medical condition. Multimedia Appendix 2 [47-58] explains the formative study by which these categories and scenarios were selected.

We used these scenarios to generate 3 case studies that highlighted the diverse ways in which health concept surveying can be used. Our first 2 case studies, which are centered around the skin cancer scenario described in the previous section, illustrate how an HIT developer or designer can use health concept surveying to decide whether to include a feature in their

HIT. Our third case study relies on all 3 scenarios to demonstrate how a human–computer interaction (HCI) researcher can use health concept surveying to elicit usable evidence without focusing on a single HIT. We restricted our investigation to a single *ActionType* (scheduling an appointment) for brevity; however, we featured multigroup path analysis in case study 3 to account for its mixed factorial design and demonstrate the expressivity of our method.

Recruitment

As our case studies were centered on health-screening apps, we recruited participants from the general population without any inclusion or exclusion criteria regarding their experiences with the relevant medical conditions. We sent calls for participation through Facebook, Reddit, and a mailing list within the University of Washington's Institute of Translational Health Sciences, a center sponsored by the National Institutes of Health's Clinical and Translational Science for connecting clinicians, patients, and other communities throughout the northwest United States. We excluded respondents who were aged <18 years or did not own a smartphone. Respondents electronically consented before viewing any of the survey materials. Respondents who completed the survey were eligible for a raffle in which 1 in 20 people would win a US $20 Amazon gift card. We used this recruitment strategy for all 3 of our case studies with approval from the University of Washington's Institutional Review Board (#00003540). Participants were restricted from taking part in multiple case studies to avoid any potential carryover effects or biases (eg, learning and fatigue).

Case Study 1: Incentivizing Clinical Visits

Overview

Our first case study investigated whether the inclusion of a wait time guarantee provides a sufficient incentive for people who would not normally seek medical attention to change their minds and get treatment. We explored this question in the context of our *serious* medical condition scenario regarding skin cancer. We recruited 54 respondents for this case study, and their demographic information can be found in Table 1.

Table 1. Demographic information for the people who completed the survey in case study 1 (N=54).

Survey demographics	Values, n (%)
Source	
Facebook	6 (11)
ITHS[a]	48 (89)
Gender	
Female	41 (76)
Male	11 (20)
Gender variant/nonconforming	2 (4)
Age (years)	
18-24	31 (57)
25-34	13 (24)
35-44	7 (13)
45-54	1 (2)
55-64	2 (4)
Smartphone operating system	
iOS	34 (63)
Android	20 (37)
Self-reported smartphone experience	
Expert or advanced	32 (59)
Intermediate	21 (39)
Novice or beginner	1 (2)

[a]ITHS: Institute of Translational Health Sciences.

Study Design

Figure 3 shows the survey design used in this study. We modified the intervention stage so that respondents were shown 1 of the 2 app descriptions at random. Half of the respondents read the SkinCheck description presented in the *Methods* section, whereas the other half saw the same description with the addition of the following text to describe a wait time incentive:

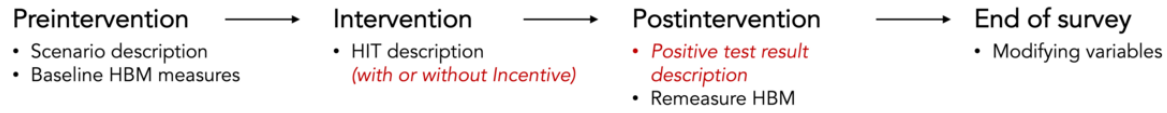
> *Because of their mobile health initiative, [Phone Company] has an exclusive partnership with dermatologists across the country. People who have a questionable mole on their skin according to*

> *SkinCheck are given a promotional code that they can redeem at their local dermatologist to guarantee a wait time no longer than 10 minutes.*

As the incentive was intended to make it easier for a person to see a clinician, we only asked respondents about how they would react to a positive test result. The study had a single-factor between-subjects design with the inclusion of an *Incentive* as the factor of interest. As our lone modifying variable, we asked respondents to rate how quickly they thought they would be able to see their physician as we hypothesized that people who did not have convenient access to a clinician would be more influenced by the incentive. We called this variable *Convenience*, and it was measured on a 7-point scale. *Incentive* and *Convenience* were connected to all major HBM constructs and outcome variables in our path diagrams.

This survey had a completion rate of 83% when we accounted for respondents who ended the survey early, satisfied the exclusion criteria, or did not correctly answer the IMC embedded in the survey. Ignoring 2 cases where respondents took more than an hour-long break while completing the survey, the median survey completion time was 8 minutes.

Figure 3. The survey structure for case study 1. The inclusion of an incentive in the health intervention technology description was randomized across respondents. HBM: Health Belief Model; HIT: health intervention technology.



TechnologyInterest

Most respondents expressed interest in using the SkinCheck app. Of the respondents who completed the survey, 54% (29/54) gave the highest rating possible for *TechnologyInterest*, 19% (10/54) gave the second-highest rating, 13% (7/54) gave the third-highest rating, and the remaining 15% (8/54) gave ratings that were either neutral or worse. The heavy bias in *TechnologyInterest* meant that a strong model fit could not be found for this outcome variable (CFI=0.839; SRMR=0.131).

ActionChangePositive

Table 2 shows the causal path coefficients for the *ActionChangePositive* model fit. Across all respondents who expressed sufficient interest in using the app, 52% (24/46) said they would not have scheduled an appointment before using the app. After being presented with a positive test result, 75% (18/24) changed their mind: 56% (10/18) were shown an incentive and 44% (8/18) were not.

Table 2. Path analysis coefficients for ActionChangePositive in case study 1 (CFI[a]=0.951; SRMR[b]=0.079).[c]

Variables	ActionChange	Seriousness	Susceptibility	Benefits	Barriers
AppResult	6.874[d]	−0.002	0.636[e]	0.024	−0.035
Incentive	1.138	−0.406	0.275	0.598	−0.361[e]
Convenience	0.128[e]	−0.492	0.055	−0.168	−0.384[f]
Seriousness	−0.005	N/A[g]	N/A	N/A	N/A
Susceptibility	0.482[e]	N/A	N/A	N/A	N/A
Benefits	0.402[e]	N/A	N/A	N/A	N/A
Barriers	−0.791[d]	N/A	N/A	N/A	N/A

[a]CFI: comparative fit index.

[b]SRMR: standardized root mean square residual.

[c]The columns indicate dependent variables, whereas the rows indicate independent variables.

[d]*P*<.001.

[e]*P*<.05.

[f]*P*<.01.

[g]N/A: not applicable.

The model fit had a large positive coefficient from *AppResult* to *ActionChangePositive* (*b*=6.874; *P*<.001), which was expected because respondents had to see a test result to change their opinion. There was also a strong positive coefficient from *AppResult* to *PerceivedSusceptibility* (*b*=0.636; *P*<.05), which supported our intuition that a positive test result should increase a person's perceived likelihood of having skin cancer. *ActionChangePositive* was heavily influenced by most of the HBM constructs. The strongest influence came from *PerceivedBarriers* (*b*=−0.791; *P*<.001), which was negative as barriers make it more difficult for a person to be able to take action.

Although there were strong coefficients from *Incentive* to all HBM constructs, the only statistically significant relationship was from *Incentive* to *PerceivedBarriers* (b=−0.361; P<.05). The fact that there is a negative coefficient between the 2 supported our expectation that the incentive would diminish the obstacles that respondents would foresee in the scenarios. Combining this finding with the strong negative coefficient from *PerceivedBarriers* to *ActionChangePositive* implies that *Incentive* had a strong positive effect on *ActionChangePositive* mediated by *PerceivedBarriers*. However, the coefficient from *Convenience* to *PerceivedBarriers* (b=−0.384; P<.01) is slightly larger in magnitude than that from *Incentive*, which indicates that the incentive was somewhat less important than the convenience of getting to a clinician in the first place. Further investigation into our data set revealed that most individuals who decided to take action after seeing a positive test result paired with an incentive gave less than a neutral rating for *Convenience*; the *Convenience* ratings for the individuals who were not shown an incentive were more evenly distributed.

Case Study 2: Presentation of Results

Overview

Our second case study investigated how the presentation of test results may influence a person's decision-making. We examined whether the inclusion of visuals that explain an algorithm's decision would engender more trust in an app's test result. As before, we explored this question in the context of our *serious* medical condition scenario regarding skin cancer. We recruited 54 respondents for this case study, and their demographic information can be found in Table 3.

Table 3. Demographic information for the people who completed the survey in case study 2 (N=54).

Survey demographics	Values, n (%)
Source	
Facebook	3 (6)
ITHS[a]	51 (94)
Gender	
Female	45 (83)
Male	8 (15)
Undisclosed	1 (2)
Age (years)	
18-24	34 (63)
25-34	13 (24)
35-44	2 (4)
45-54	2 (4)
55-64	2 (4)
Undisclosed	1 (2)
Smartphone operating system	
iOS	39 (72)
Android	15 (28)
Self-reported smartphone experience	
Expert or advanced	28 (52)
Intermediate	26 (48)

[a]ITHS: Institute of Translational Health Sciences.

Study Design

Figure 4 shows the survey design used in this study. We modified the postintervention stage so that respondents would be asked to react to both positive and negative test results. Instead of explaining the test result in a paragraph, as in the previous case study, respondents were shown 1 of 2 result screen concepts, illustrated in Figure 5 [47], at random. Both screens were derived from the DermoScreen app by Wadhawan et al [47], which explains diagnostic decisions using the ABCD rule of dermatoscopy [59].

The study had a single-factor between-subjects design with the inclusion of *Visuals* as the factor of interest. As our lone modifying variable, we asked respondents about their highest level of education as we hypothesized that reading comprehension would affect their understanding of the visualizations; we called this variable *Education*. *Visuals* and *Education* were connected to all major HBM constructs and outcome variables in our path diagrams.

This survey had a completion rate of 82% when we accounted for respondents who ended the survey early, satisfied the exclusion criteria, or did not correctly answer the IMC

embedded in the survey. Ignoring 1 case when a respondent took more than an hour-long break while completing the survey, the median survey completion time was 9 minutes.

Figure 4. The 2 possible interface options that respondents could have been shown in case study 2 when presented with a positive test result: the interface with text descriptions only (left) and the interface with text and visuals to illustrate how the results were obtained (right). The interfaces were primarily inspired by the DermoScreen app by Wadhawan et al [47]. HBM: Health Belief Model; HIT: health intervention technology.

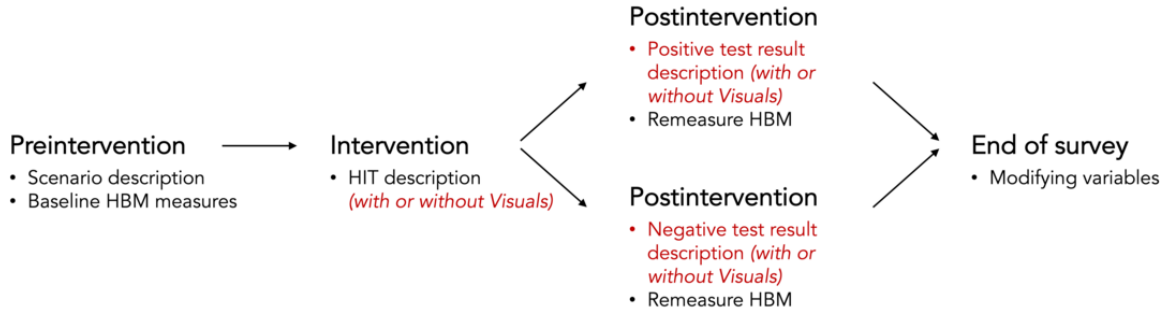
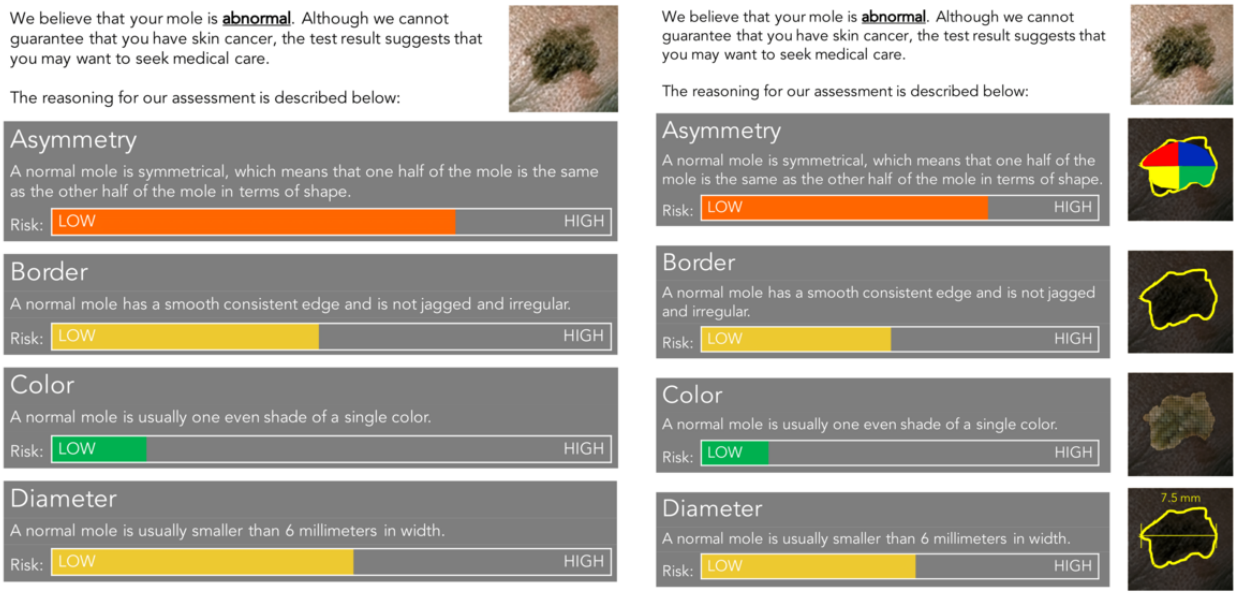


Figure 5. The two possible interface options that respondents could have been shown in Case Study 2 when presented with a positive test result: (left) the interface with text descriptions only and (right) the interface with text and visuals to illustrate how the results were obtained. The interfaces were primarily inspired by Wadhawan et al.'s [47] DermoScreen app.



TechnologyInterest

Most respondents expressed interest in using the SkinCheck app. Of the respondents who completed the survey, 56% (30/54) gave the highest rating possible for *TechnologyInterest*, 19% (10/54) gave the second-highest rating, 15% (8/54) gave the third-highest rating, and the remaining 11% (6/54) gave ratings that were either neutral or worse. The heavy bias in *TechnologyInterest* meant that a strong model fit could not be found for this outcome variable (CFI=0.874; SRMR=0.096).

ActionChangePositive

Across all respondents who expressed sufficient interest in using the app, there were 56% (27/48) of cases when people said that they would not have acted before using the app. After being presented with a positive test result, 78% (21/27) changed their mind: 38% (8/21) were shown visuals and 62% (13/21) were not. The inclusion of explanations clearly had an impact on people's reaction to the positive test result as the frequency of *ActionChangePositive* was much higher than in the first case study. In fact, there were so few cases when people did not act even after seeing a positive test result that there was not enough data to generate a meaningful model fit (CFI=0.644; SRMR=0.197).

ActionChangeNegative

Table 4 shows the causal path coefficients for the *ActionChangeNegative* model fit. Across all respondents who expressed sufficient interest in using the app, there were 52% (25/48) of cases when people said that they would have acted before using the app. After being presented with a negative test result, 48% (12/25) changed their mind: 50% (6/12) were shown visuals and 50% (6/12) were not.

As *ActionChangeNegative* is positive when a person is swayed to not act in the postintervention stage, we expected many of the path coefficients to be negated relative to those observed with *ActionChangePositive* in the first case study. This expectation was confirmed in a couple of instances. First, the negative coefficient from *AppResult* to *PerceivedSusceptibility*

(b=−0.222; P<.05) confirmed our intuition that a negative test result should decrease a person's belief that they had skin cancer in this scenario. Second, the negative coefficient from *PerceivedSeriousness* to *ActionChangeNegative* (b=−0.220; P<.05) showed that people who were not as concerned about skin cancer were more likely to change their course of action.

As we hypothesized, including additional information in the form of visuals strengthened respondents' confidence in their test results. This was reflected in the negative coefficient from *Visuals* to *PerceivedSusceptibility* (b=−0.961; P<.01); when shown a negative test result with visuals, respondents were less likely to believe they had skin cancer. However, *PerceivedSusceptibility* was not influential on *ActionChangeNegative* (b=−0.056, not significant); therefore, the inclusion of visuals had neither a direct nor indirect effect on a person's decision to change their action. We also found that *Education* was not an influential factor for any of the measured constructs or outcome variables.

Table 4. Path analysis coefficients for ActionChangeNegative in case study 2 (CFI[a]=0.961; SRMR[b]=0.078).[c]

Variables	ActionChange	Seriousness	Susceptibility	Benefits	Barriers
AppResult	6.588[d]	0.000	−0.222[e]	0.000	−0.342
Visuals	0.231	0.235	−0.961[f]	−0.591	−0.610
Education	0.087	0.086	0.037	0.107	0.055
Seriousness	−0.220[e]	N/A[g]	N/A	N/A	N/A
Susceptibility	−0.056	N/A	N/A	N/A	N/A
Benefits	−0.233[e]	N/A	N/A	N/A	N/A
Barriers	0.223	N/A	N/A	N/A	N/A

[a]CFI: comparative fit index.

[b]SRMR: standardized root mean square residual.

[c]The columns indicate dependent variables, whereas the rows indicate independent variables.

[d]P<.001.

[e]P<.05.

[f]P<.01.

[g]N/A: not applicable.

Case Study 3: Accuracy

Overview

In our third and final case study, we explored the trade-off between false positives and false negatives across medical conditions of varying concern and severity. We leveraged all three of our scenarios (*common*, *serious*, and *stigmatizing*) in a mixed factorial study design, thus necessitating more participants. In total, 263 respondents completed the survey from start to finish, and their demographic information can be found in Table 5.

Table 5. Demographic information for the people who completed the survey in case study 3 (N=263).

Survey demographics	Values, n (%)
Source	
Facebook	16 (6.1)
ITHS[a]	240 (91.3)
Reddit	3 (1.1)
Other	4 (1.5)
Gender	
Female	202 (76.8)
Male	45 (17.1)
Transgender male	5 (1.9)
Gender variant/nonconforming	7 (2.7)
Self-identify	1 (0.4)
Undisclosed	3 (1.1)
Age (years)	
18-24	145 (55.1)
25-34	84 (32)
35-44	17 (6.5)
45-54	8 (3.1)
55-64	3 (1.1)
≥65	3 (1.1)
Undisclosed	3 (1.1)
Smartphone operating system	
iOS	170 (64.6)
Android	93 (35.4)
Self-reported smartphone experience	
Expert or advanced	146 (55.5)
Intermediate	115 (43.7)
Novice or beginner	2 (0.8)

[a]ITHS: Institute of Translational Health Sciences.

Study Design

Figure 6 shows the survey design for this study, which required changes in both the intervention and postintervention stages. The app descriptions included information about their classification sensitivity and specificity; sensitivity refers to the proportion of people who are correctly identified as having the medical condition out of all those who have it, whereas specificity refers to the proportion of people who are correctly identified as not having the medical condition out of all those who do not have it. Because the general public is more adept at reasoning about counts than fractional quantities [60], the sensitivity and specificity rates were presented with counts and icon arrays. An example of the accompanying text is provided as follows:

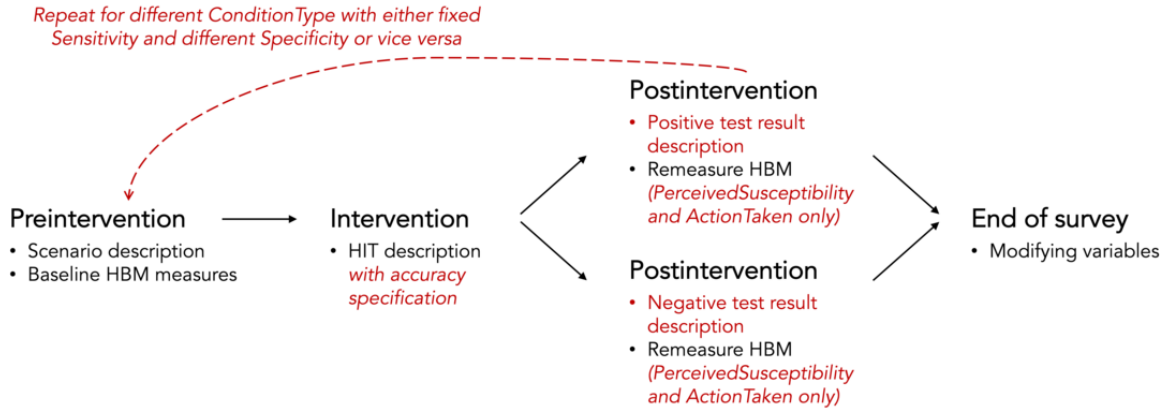
> *Out of every 100 people who have a sinus infection, SinusCheck correctly told 65 people that they had a sinus infection.*
>
> *Out of every 100 people who do not have a sinus infection, SinusCheck correctly told 80 people that they did not have a sinus infection.*

The survey was used in a 3×3×3 mixed factorial study design. Each respondent read all three scenarios—pink eye (*common*), skin cancer (*serious*), and halitosis (*stigmatizing*)—making *ConditionType* a within-subjects factor. The presentation order of the scenarios was counterbalanced across all subjects. A total of 3 equally spaced levels of sensitivity and specificity were investigated—65%, 80%, and 95%—producing 9 possible combinations that described the overall accuracy of the apps. Each app for each respondent was assigned 1 of the 9 combinations at random, making *Sensitivity* and *Specificity* between-subjects factors. Although there is an inherent trade-off

between sensitivity and specificity when the underlying classification algorithm is fixed, we treated them as independent variables in our study design and analyses. As respondents had to go through multiple scenarios, we shortened the survey by only remeasuring *PerceivedSusceptibility* and *ActionTaken* during the postintervention stage. The other major HBM constructs were not remeasured as we assumed that they should not be influenced by app accuracy. As such, *Sensitivity* and *Specificity* were connected to *PerceivedSusceptibility* and the outcome variables in our path diagrams, and *ConditionType* was used as the grouping variable for multigroup path analysis.

This survey had a completion rate of 73% when we accounted for respondents who ended the survey early, satisfied the exclusion criteria, or did not correctly answer the IMCs embedded in the survey. Ignoring 14 cases when respondents took more than an hour-long break while completing the survey, the median survey completion time was 16 minutes.

Figure 6. The survey structure for case study 3. Respondents were shown 3 different health intervention technologies (HITs)—1 for each ConditionType. The 3 HITs either had the same sensitivity and varied in specificity or had the same specificity and varied in sensitivity. Respondents were asked to react to positive and negative app results in a randomized order. Only PerceivedSusceptibility and ActionChange were remeasured in the postintervention stages to shorten the survey length. HBM: Health Belief Model; HIT: health intervention technology.



TechnologyInterest

Table 6 shows the causal path coefficients for the *TechnologyInterest* model fit. The path coefficients from *Sensitivity* and *Specificity* to *TechnologyInterest* were sizable and positive across all scenarios, confirming that higher accuracy made the apps more attractive. In fact, the effect was so strong that those coefficients were larger and more statistically significant than those from the HBM constructs. This suggests that respondents were willing to use these apps regardless of their perception of the medical conditions' threat as long as they knew that the app was accurate.

Table 6. Path analysis coefficients for ActionChangeNegative in case study 3 (CFI[a]=0.956; SRMR[b]=0.075).[c]

Variables	Technology Interest		
	Common	Serious	Stigmatizing
Seriousness	−0.120	0.101	0.129
Susceptibility	0.206[d]	0.120[e]	0.104
Sensitivity	0.416[f]	0.357[f]	0.268[d]
Specificity	0.461[f]	0.300[d]	0.292[f]

[a]CFI: comparative fit index.

[b]SRMR: standardized root mean square residual.

[c]The columns indicate dependent variables, whereas the rows indicate independent variables.

[d]*P*<.01.

[e]*P*<.05.

[f]*P*<.001.

Overall accuracy was most valued for the *common* condition (*Sensitivity*: *b*=0.416, *P*<.001; *Specificity*: *b*=0.461, *P*<.001), followed by the *serious* (*Sensitivity*: *b*=0.357, *P*<.001; *Specificity*: *b*=0.300, *P*<.01) and *stigmatizing* (*Sensitivity*: *b*=0.268, *P*<.01; *Specificity*: *b*=0.292, *P*<.001) conditions. Respondents preferred apps with higher accuracy; however, they attributed more importance to sensitivity or specificity depending on the scenario; they placed more importance on specificity for the *common* and *stigmatizing* conditions, whereas they placed more importance on sensitivity for the *serious* condition. Sensitivity and specificity were treated independently in our analysis; therefore, these results do not account for the fact that improving one metric often requires compromising the other during the development of the classification model.

Nevertheless, this result suggests that respondents had an inherent knowledge about the notion of prevalence and how it relates to diagnostic decision-making. *Common* and *stigmatizing* conditions are typically prevalent; therefore, prioritizing specificity may indicate that respondents were eager to use an app's test result to *rule out* having the condition. *Serious* conditions are often less prevalent; therefore, prioritizing sensitivity may indicate that respondents were eager to *rule in* having the condition.

ActionChangePositive

Table 7 shows the causal path coefficients for the *ActionChangePositive* model fit. Across all respondents who expressed sufficient interest in using any of the 3 apps, there were 56.5% (359/635) of cases when respondents said that they would not have taken action before using the app. After being presented with a positive test result, 46.2% (166/359) changed their mind: 28.3% (47/166) in the *common* scenario, 42.2% (70/166) in the *serious* scenario, and 29.5% (49/166) in the *stigmatizing* scenario.

Table 7. Path analysis coefficients for ActionChangePositive in case study 3 (CFI[a]=0.981; SRMR[b]=0.078).[c]

Variables	Common		Serious		Stigmatizing	
	ActionChange	Susceptibility	ActionChange	Susceptibility	ActionChange	Susceptibility
AppResult	6.962[d]	0.398[d]	6.521[d]	1.518[d]	5.900[d]	0.474[e]
Sensitivity	−0.262	0.283[e]	0.011	0.014	−0.004	0.204[f]
Specificity	−0.095	0.095	−0.212	−0.093	0.033	−0.032
Seriousness	−0.149	N/A[g]	−0.124	N/A	0.168	N/A
Susceptibility	0.426[e]	N/A	0.509[d]	N/A	0.474[d]	N/A
Benefits	0.226[e]	N/A	0.273	N/A	0.055	N/A
Barriers	−0.137	N/A	−0.061	N/A	−0.038	N/A

[a]CFI: comparative fit index.

[b]SRMR: standardized root mean square residual.

[c]The columns indicate dependent variables, whereas the rows indicate independent variables.

[d]$P<.001$.

[e]$P<.01$.

[f]$P<.05$.

[g]N/A: not applicable.

Although there were large positive coefficients from *AppResult* to *ActionChangePositive* and *PerceivedSusceptibility* across all scenarios, the magnitude and significance of those coefficients varied across the medical conditions. The coefficient from *AppResult* to *PerceivedSusceptibility* for the *Serious* condition (b=1.518; $P<.001$) was 3 times as large and more significant than the corresponding coefficients for the *common* (b=0.398; $P<.001$) and *stigmatizing* (b=0.474; $P<.01$) conditions. Again, this result suggests that respondents may have been eager to use the positive test result from an app to *rule in* having a *serious* condition.

Sensitivity had a significant positive effect on *PerceivedSusceptibility* for the *common* (b=0.283; $P<.001$) and *stigmatizing* (b=0.204; $P<.05$) conditions. There were also significant positive coefficients between *PerceivedSusceptibility* and *ActionChangePositive* in those scenarios (common: b=0.426, $P<.01$; stigmatizing: b=0.474, $P<.001$), which means that *Sensitivity* had a strong effect on *ActionChangePositive* mediated by *PerceivedSusceptibility* in those scenarios. In other words, respondents were more likely to be convinced to change their course of action after seeing a positive test result when the app had a higher sensitivity. *Sensitivity* did not have a significant effect on *PerceivedSusceptibility* for the *serious* condition (b=0.014, not significant), implying that respondents were equally willing to accept a positive test result across the presented sensitivity rates in that scenario. *Specificity* did not have a statistically significant effect on either *ActionChangePositive* or *PerceivedSusceptibility* for any of the scenarios. Although sensitivity corresponds to a test's true negative rate, this finding is still notable as sensitivity affects how a positive test result should be interpreted according to Bayesian statistics.

ActionChangeNegative

Table 8 shows the causal path coefficients for the *ActionChangeNegative* model fit. We note that this model had borderline significance according to our fit statistics, satisfying the threshold for SRMR but not for CFI. Across all respondents who expressed sufficient interest in using any of the 3 apps, there were 41.9% (266/635) of cases when respondents said that they would have taken action before using the app. After being presented with a negative test result, 51.9% (138/266) changed their minds: 39.1% (54/138) in the *common* scenario, 37.7% (52/138) in the *serious* scenario, and 23.2% (32/138) in the *stigmatizing* scenario.

Table 8. Path analysis coefficients for ActionChangeNegative in case study 3 (CFI[a]=0.925; SRMR[b]=0.077).[c]

Variables	Common		Serious		Stigmatizing	
	ActionChange	Susceptibility	ActionChange	Susceptibility	ActionChange	Susceptibility
AppResult	6.230[d]	−1.999[d]	6.144[d]	−0.970[d]	6.833[d]	−2.191[d]
Sensitivity	0.022	−0.196	−0.022	−0.026	0.003	0.056
Specificity	−0.103[e]	0.094	0.063	−0.185[e]	−0.192	−0.119
Seriousness	−0.451[e]	N/A[f]	−0.182	N/A	−0.169	N/A
Susceptibility	−0.311[d]	N/A	−0.358[d]	N/A	−0.212[d]	N/A
Benefits	0.001	N/A	0.125	N/A	−0.196	N/A
Barriers	0.105	N/A	0.083	N/A	−0.176	N/A

[a]CFI: comparative fit index.

[b]SRMR: standardized root mean square residual.

[c]The columns indicate dependent variables, whereas the rows indicate independent variables.

[d]$P<.001$.

[e]$P<.05$.

[f]N/A: not applicable.

As with the model fit for *ActionChangePositive*, there were statistically significant coefficients from *AppResult* to *PerceivedSusceptibility* and *ActionChangeNegative*; however, their magnitude varied across *ConditionType*. The coefficients from *AppResult* to *PerceivedSusceptibility* for the *common* ($b=-1.999$; $P<.001$) and *stigmatizing* ($b=-2.191$; $P<.001$) conditions were nearly double the corresponding coefficient for the *serious* condition ($b=-0.970$; $P<.001$), suggesting that respondents may have been eager to use the negative test result from those apps to *rule out* having those conditions.

In the *serious* condition scenario, significant negative coefficients were found from *specificity* to *PerceivedSusceptibility* ($b=-0.185$; $P<.001$) and *PerceivedSusceptibility* to *ActionChangeNegative* ($b=-0.358$; $P<.001$). This combination of results implies that respondents were more likely to be convinced to change their course of action after seeing a negative test result when the app had a higher specificity. *Specificity* did not have a significant effect on *PerceivedSusceptibility* in either the *common* ($b=0.094$, not significant) or the *stigmatizing* ($b=-0.119$, not significant) conditions, indicating that respondents were equally willing to accept a negative test result across the presented specificity rates in those scenarios. *Sensitivity* did not have a statistically significant effect on either *ActionChangeNegative* or *PerceivedSusceptibility* for any of the scenarios, which mirrors the earlier findings with respect to *specificity* and positive test results.

Discussion

Principal Findings

We sought to develop a low-burden method for projecting the adoption and acceptability of an HIT, given different design variations. Our contribution toward this goal—the health concept surveying method—supports HIT investigators in advancing their own HITs while generating usable evidence for the broader research community. Our 3 case studies highlight the different types of actionable feedback and usable evidence that can be elicited using our survey instrument without deploying a working HIT prototype.

Our first case study showed that a wait time incentive might support some individuals in overcoming barriers that could prevent them from visiting a dermatologist. However, many participants said that they would be persuaded to act without an incentive. This result suggests that HIT developers in this scenario may want to consider additional messaging that targets other facets of the HBM, such as the perceived susceptibility people have to skin cancer or the perceived benefits of seeking a second opinion. We also found that access to convenient health care was an important factor in people's decision-making; therefore, developers in this scenario may want to examine whether this is an important issue to address for their target audience.

Our second case study showed that SkinCheck's baseline explanation could be convincing enough to sway a person to visit a clinician when they received a positive test result. The inclusion of visuals increased individuals' trust in negative test results; however, this was not enough to significantly affect people's decision-making. In fact, we found that the main driving factor for people who decided not to act after seeing a negative test result was the perceived seriousness of skin cancer. This presents an interesting challenge for HIT designers. Lowering a person's concern about the severity of a medical condition could have major consequences, including the fact that they may ignore a positive test result later on because of their newfound understanding of the condition. Instead, HIT designers in this scenario may want to consider using a language that diminishes a person's short-term concerns but encourages repeated testing in the near future.

Our third case study suggests that researchers may want to consider the trade-off between sensitivity and specificity in the context of their target medical condition. Kay et al [61] elicited

similar findings through a survey instrument they created to understand the acceptability of precision and recall across various sensor-based technologies. In an example involving a home alarm system, they showed that participants were more willing to accept false alarms when the system had a benign intervention (eg, contacting the homeowners via SMS text message) than when the system had an intrusive intervention (eg, automatically alerting the police). To improve the user experience that people have with a classifier-based application, HIT developers may consider adjusting the final decision threshold of their classifier to minimize errors that people are more prone to believe. However, doing so may serve as an expedient solution to the greater challenge of helping ordinary people with Bayesian reasoning.

Other Design Decisions for Exploration

We explored the influence of 3 different design choices on outcomes relevant to HITs (incentives, results presentation, and accuracy trade-offs); however, there are many others that would be interesting to explore in future work. One of those factors would be the HIT's price. When we first piloted our studies, we stated that the apps could be purchased on app stores for US $0.99. We selected such a low cost as we were worried that a free app would appear illegitimate; however, an expensive app would diminish interest to the point that we would not receive feedback from respondents. However, some of the respondents in our pilot study felt that a US $0.99 app appeared *less legitimate than a free app* and *cheap*, so we instead crafted scenarios in which the app was already included on the respondents' phones. The economics research community has debated the relationship between price and perceived product quality; some researchers argue that there is generally a positive correlation between price and quality [62], whereas others argue that the 2 are only correlated under contrived scenarios [63].

Another factor that influences the perceived quality of technology is endorsements [64]. App stores, smartphone manufacturers, special interest groups, and physicians can all endorse technologies, serving as a *seal of approval* that may imbue an HIT with legitimacy. A limitation of our survey instrument is that it is difficult to convey an endorsement to respondents without explicitly drawing the respondents' attention to it. Endorsements can appear in many places—commercials, supplemental materials, or websites—that may not be as conspicuous as mentioning would be done in the survey. Determining a more natural way of introducing endorsements within health concept surveying could be a potential avenue for future work.

Alternative and Complementary Approaches

Health concept surveying is one of many early-stage quantitative research methods that developers and designers can use to further their understanding of HITs. Conjoint analysis and discrete choice experiments elicit preferences by asking participants to pick between options with 1 or many feature variations in a head-to-head comparison [65]. Another relevant technique is judgment analysis [66,67], where feature preferences are gathered by comparing the decisions that participants make in hypothetical scenarios against a predefined oracle or reference group. All of these methods have been used to investigate people's decision-making in the health domain [68-70]; however, health concept surveying has the advantage of being designed so that investigators can project both the adoption of an HIT and the acceptability of an HIT's suggestions. By accounting for intrinsic and extrinsic factors that can influence these distal outcomes, health concept surveying is able to elicit usable evidence that HIT developers and designers can apply to their own HITs.

We view health concept surveying as being complementary to qualitative research methods such as focus group interviews, which give participants the chance to verbalize their thoughts and decision-making in a richer way than what can be gathered through a survey. That said, health concept surveying is far more efficient to scale. Focus groups must be run with 5 to 10 participants at a time, and investigators must often conduct multiple sessions to reach diverse populations or gather feedback on new design iterations. Each new session incurs an additional time investment for both the interviews and the qualitative analyses, making focus groups difficult to scale as an HIT evolves. In addition, focus groups have known confounds such as group-think or dominance by 1 or 2 individuals, even in light of techniques to mitigate these confounds [71]. With health concept surveying, adding more participants simply requires distributing the survey to more people and then rerunning the same analysis code as before, imposing no additional burden beyond what is required for recruitment. Health concept surveying also helps investigators systematically analyze the influence of all the variables involved in people's decision-making, which can otherwise be difficult for participants to articulate and for investigators to translate into usable evidence. We hope that our work inspires HCI researchers to explore how people can incorporate psychological frameworks into other evaluation techniques.

Limitations

Several psychological frameworks for explaining behavior rely on the belief that intention is a strong predictor of behavior. The correlation between intention and behavior has been supported by research on health-related topics such as dieting [72], physical activity [73,74], and weight loss [75]. Nevertheless, people's behavioral intentions or expected actions do not always lead to completing the action because of the emergence of unforeseen barriers or changing beliefs over time. Psychologists have called this phenomenon *the intention–behavior gap* [76,77]. This potential disconnect exists in most early-stage evaluation methods; however, the gap may be particularly relevant to health concept surveying as intention in scenario-based study designs may not translate to real-world actions, and there are no consequences to hypothetical decisions. Despite these shortcomings, there are steps that HIT investigators can take to engender more confidence in their survey responses. We recruited respondents from the public; however, developers and designers who are creating an HIT for a specific audience may want to recruit participants who are either in an at-risk demographic or actively seeking solutions in the HIT's target domain. As realism is an important mediator in the intention–behavior gap, we also suggest that investigators craft their scenarios with the help of domain experts to make the scenarios as realistic as possible. Investigators could even add

questions to their surveys that measure the degree to which respondents resonate with their scenarios; such measures could be used to either filter responses or create an additional modifying variable in the analyses.

HIT investigators may also want to consider focusing on short-term actions rather than long-term goals (eg, *I intend to eat more vegetables for dinner today* vs *I intend to lose 10 pounds this month*) when querying how a person would respond to an HIT; intention is believed to be a weaker predictor for long-term goals as completing them requires more self-efficacy and coordination to complete [76]. Finally, the health action process approach of Schwarzer [78] separates preintentional motivation and postintentional volition when measuring the likelihood of action; therefore, doing the same in health concept surveying may be beneficial.

To ensure that we were collecting meaningful responses, we also had to create plausible scenarios. We validated the scenarios used in our work through a pilot study using an abridged version of our survey instrument. Researchers who are investigating high-level questions as we did in our third case study would want to repeat this procedure; however, an HIT developer interested in advancing a particular HIT design while generating usable evidence may only need to assess scenario plausibility. We used a single question that explicitly asked respondents how plausible they believed a scenario to be; however, future researchers may want to investigate the nuances of plausibility through multiple questions. A person may believe a scenario is plausible as the health issue in question is common for their demographic or to people who engage in similar behaviors, or they may believe it is plausible because they do not have enough knowledge about the issue to know better. Researchers interested in examining HIT design decisions across multiple scenarios may also want to consider making their scenarios publicly available for future use. Sharing a common set of prevalidated scenarios would standardize the context of findings related to the same topic (eg, physical activity, step counting, and exercise).

Conclusions

As more HITs transition from research to practice, it is important for HCI researchers to examine how those technologies will be received by the general population. Although one-off user studies provide actionable feedback for a specific HIT, they rarely provide insights that benefit other HIT creators. Our method, health concept surveying, attempts to strike a balance between actionable feedback and usable evidence. Using the HBM, health concept surveying disentangles proximal cognitive factors from HIT design decisions to explain *how* and *why* certain features are preferred. We used health concept surveying in 3 case studies to demonstrate the range of questions it can support and discussed the implications of the findings in each case. We hope that researchers will continue using health concept surveying in the future to better our understanding of HITs and accelerate their development.

Acknowledgments

This project was supported by funding from the National Science Foundation (NSF IIS-1813675, IIS-1553167) and the National Institutes of Health (NIH R21 DK117431 and NIH R01 LM012810). The authors would like to thank Wanda Pratt for her guidance during the early stages of this research. The authors would also like to thank Gary Hsieh for his feedback on the framing of this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
An example of the health concept surveying survey instrument was applied to a design concept called SkinCheck.
[DOCX File , 75 KB - humanfactors_v9i1e30474_app1.docx]

Multimedia Appendix 2
A preliminary study was conducted to select the scenarios and design concepts that appear in this paper.
[DOCX File , 459 KB - humanfactors_v9i1e30474_app2.docx]

References

1. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. J Med Internet Res 2020 Dec 09;22(12):e21687 [FREE Full text] [doi: 10.2196/21687] [Medline: 33295292]
2. Kanthawala S, Joo E, Kononova A, Peng W, Cotten S. Folk theorizing the quality and credibility of health apps. Mobile Media Commun 2018 Oct 02;7(2):175-194. [doi: 10.1177/2050157918796859]
3. Choe E, Lee B, Munson S, Pratt W, Kientz J. Persuasive performance feedback: the effect of framing on self-efficacy. AMIA Annu Symp Proc 2013;2013:825-833 [FREE Full text] [Medline: 24551378]

4. Coyne I, Prizeman G, Sheehan A, Malone H, While AE. An e-health intervention to support the transition of young people with long-term illnesses to adult healthcare services: design and early use. Patient Educ Couns 2016 Dec;99(9):1496-1504. [doi: 10.1016/j.pec.2016.06.005] [Medline: 27372524]
5. Wagner A, Boulton M, Sun X, Mukherjee B, Huang Z, Harmsen I, et al. Perceptions of measles, pneumonia, and meningitis vaccines among caregivers in Shanghai, China, and the health belief model: a cross-sectional study. BMC Pediatr 2017 Jun 12;17(1):143 [FREE Full text] [doi: 10.1186/s12887-017-0900-2] [Medline: 28606106]
6. Champion VL. The relationship of breast self-examination to health belief model variables. Res Nurs Health 1987 Dec;10(6):375-382. [doi: 10.1002/nur.4770100605] [Medline: 3423309]
7. Collins LM, Dziak JJ, Li R. Design of experiments with multiple independent variables: a resource management perspective on complete and reduced factorial designs. Psychol Methods 2009 Sep;14(3):202-224 [FREE Full text] [doi: 10.1037/a0015826] [Medline: 19719358]
8. Collins LM, Dziak JJ, Kugler KC, Trail JB. Factorial experiments: efficient tools for evaluation of intervention components. Am J Prev Med 2014 Oct;47(4):498-504 [FREE Full text] [doi: 10.1016/j.amepre.2014.06.021] [Medline: 25092122]
9. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34S(Suppl):1220-1228 [FREE Full text] [doi: 10.1037/hea0000305] [Medline: 26651463]
10. Liao P, Klasnja P, Tewari A, Murphy SA. Sample size calculations for micro-randomized trials in mHealth. Stat Med 2016 May 30;35(12):1944-1971 [FREE Full text] [doi: 10.1002/sim.6847] [Medline: 26707831]
11. Dallery J, Raiff BR. Optimizing behavioral health interventions with single-case designs: from development to dissemination. Transl Behav Med 2014 Sep 7;4(3):290-303 [FREE Full text] [doi: 10.1007/s13142-014-0258-z] [Medline: 25264468]
12. Dallery J, Cassidy RN, Raiff BR. Single-case experimental designs to evaluate novel technology-based health interventions. J Med Internet Res 2013 Feb 08;15(2):e22 [FREE Full text] [doi: 10.2196/jmir.2227] [Medline: 23399668]
13. Norman D, Draper S. User Centered System Design; New Perspectives on Human-Computer Interaction. USA: Erlbaum Associates; 1986.
14. Xu X, Li H, Shan S. Understanding the health behavior decision-making process with situational theory of problem solving in online health communities: the effects of health beliefs, message credibility, and communication behaviors on health behavioral intention. Int J Environ Res Public Health 2021 Apr 23;18(9):4488 [FREE Full text] [doi: 10.3390/ijerph18094488] [Medline: 33922583]
15. Christoforakos L, Diefenbach S. Idealization effects in UX evaluation at early concept stages: challenges of low-fidelity prototyping. In: Advances in Usability, User Experience and Assistive Technology. Cham: Springer; 2019:3-14.
16. Tittle C, Hill R. Attitude measurement and prediction of behavior: an evaluation of conditions and measurement techniques. In: Sociological Methods. Oxfordshire, United Kingdom: Routledge; 2018.
17. Klasnja P, Hekler E, Korinek EV, Harlow J, Mishra S. Toward usable evidence: optimizing knowledge accumulation in HCI research on health behavior change. Proc SIGCHI Conf Hum Factor Comput Syst 2017 May;2017:3071-3082 [FREE Full text] [doi: 10.1145/3025453.3026013] [Medline: 30272059]
18. Lewis CC, Klasnja P, Powell BJ, Lyon AR, Tuzzio L, Jones S, et al. From classification to causality: advancing understanding of mechanisms of change in implementation science. Front Public Health 2018 May 7;6:136 [FREE Full text] [doi: 10.3389/fpubh.2018.00136] [Medline: 29868544]
19. Bachmann L, Mühleisen A, Bock A, Riet GT, Held U, Kessels A. Vignette studies of medical choice and judgement to study caregivers' medical decision behaviour: systematic review. BMC Med Res Methodol 2008 Jul 30;8:50 [FREE Full text] [doi: 10.1186/1471-2288-8-50] [Medline: 18664302]
20. Evans S, Roberts M, Keeley J, Blossom J, Amaro C, Garcia A, et al. Vignette methodologies for studying clinicians' decision-making: validity, utility, and application in ICD-11 field studies. Int J Clin Health Psychol 2015;15(2):160-170 [FREE Full text] [doi: 10.1016/j.ijchp.2014.12.001] [Medline: 30487833]
21. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar 19;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
22. Hermes ED, Lyon AR, Schueller SM, Glass JE. Measuring the implementation of behavioral intervention technologies: recharacterization of established outcomes. J Med Internet Res 2019 Jan 25;21(1):e11752. [doi: 10.2196/11752] [Medline: 30681966]
23. Hochbaum GM. Public Participation in Medical Screening Programs: A Socio-Psychological Study. Washington, DC: US Department of Health, Education, and Welfare, Public Health Service, Bureau of State Services, Division of Special Health Services, Tuberculosis Program; 1958.
24. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
25. Triantafyllidis AK, Velardo C, Salvi D, Shah SA, Koutkias VG, Tarassenko L. A survey of mobile phone sensing, self-reporting, and social sharing for pervasive healthcare. IEEE J Biomed Health Inform 2017 Jan;21(1):218-227. [doi: 10.1109/JBHI.2015.2483902] [Medline: 26441432]

26. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2011; Vancouver BC Canada p. 3063-3072. [doi: 10.1145/1978942.1979396]
27. Hekler E, Klasnja P, Froehlich J, Buman M. Mind the theoretical gap: interpreting, using, and developing behavioral theory in HCI research. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing System; April 27 - May 2, 2013; Paris France p. 3307-3316. [doi: 10.1145/2470654.2466452]
28. Fogg BJ. Persuasive technology: using computers to change what we think and do. Ubiquity 2002 Dec;2002(December):2. [doi: 10.1145/764008.763957]
29. Mullen PD, Hersey JC, Iverson DC. Health behavior models compared. Soc Sci Med 1987;24(11):973-981. [doi: 10.1016/0277-9536(87)90291-7] [Medline: 3616691]
30. Corace KM, Srigley JA, Hargadon DP, Yu D, MacDonald TK, Fabrigar LR, et al. Using behavior change frameworks to improve healthcare worker influenza vaccination rates: a systematic review. Vaccine 2016 Jun 14;34(28):3235-3242 [FREE Full text] [doi: 10.1016/j.vaccine.2016.04.071] [Medline: 27155491]
31. Lewis MS. Health behavior change models to prevent noise-induced auditory dysfunction. Hear J 2020;73(10):36-38. [doi: 10.1097/01.HJ.0000719800.96632.d5]
32. Sutton S. Health behavior: psychosocial theories. In: International Encyclopedia of the Social & Behavioral Sciences. Amsterdam: Elsevier; 2001:6499-6506.
33. Fishbein M. Attitude and the prediction of behavior. In: Readings in Attitude Theory and Measurement. Hoboken, New Jersey, United States: John Wiley & Sons Inc; 1967:477-492.
34. Prochaska J, DiClemente C. The Transtheoretical Approach: Crossing Traditional Boundaries of Therapy. New York: Dow Jones-Irwin; 1984.
35. Taylor D, Bury M, Campling N, Carter S, Garfied S, Newbould J, et al. A Review of the use of the Health Belief Model (HBM), the Theory of Reasoned Action (TRA), the Theory of Planned Behaviour (TPB) and the Trans-Theoretical Model (TTM) to study and predict health related behaviour change. National Institute for Health and Clinical Excellence. 2006. URL: https://tinyurl.com/2p9e2s2x [accessed 2019-05-02]
36. Ajzen I. Models of human social behavior and their application to health psychology. Psychol Health 1998 Jul;13(4):735-739. [doi: 10.1080/08870449808407426]
37. Armitage CJ, Conner M. Social cognition models and health behaviour: a structured review. Psychol Health 2000 Mar;15(2):173-189. [doi: 10.1080/08870440008400299]
38. Harrison J, Mullen P, Green LW. A meta-analysis of studies of the Health Belief Model with adults. Health Educ Res 1992 Mar;7(1):107-116 [FREE Full text] [doi: 10.1093/her/7.1.107] [Medline: 10148735]
39. Urich A. Methods for stress management. Creative Commons Attribution ShareAlike. 2017. URL: https://psu.pb.unizin.org/kines082/ [accessed 2021-12-16]
40. Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation checks: detecting satisficing to increase statistical power. J Experiment Soc Psychol 2009 Jul;45(4):867-872. [doi: 10.1016/j.jesp.2009.03.009]
41. Jöreskog KG. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika 1969 Jun;34(2):183-202. [doi: 10.1007/bf02289343]
42. Chin W. Commentary: issues and opinion on structural equation modeling. MIS Q 1998;22(1):1-8 [FREE Full text]
43. Kline R. Principles and Practice of Structural Equation Modeling. New York: Guilford Press; 2005.
44. Hooper D, Coughlan J, Mullen M. Structural equation modelling: guidelines for determining model fit. Electron J Bus Res Methods 2008;6(1):53-60 [FREE Full text]
45. Moss S. Fit indices for structural equation modeling. Sicotests. 2016. URL: https://www.sicotests.com/psyarticle.asp?id=277 [accessed 2019-03-26]
46. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
47. Wadhawan T, Situ N, Rui H, Lancaster K, Yuan X, Zouridakis G. Implementation of the 7-point checklist for melanoma detection on smart handheld devices. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2011 Presented at: Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 30 - Sept 3, 2011; Boston, MA, USA. [doi: 10.1109/iembs.2011.6090866]
48. Chandra RK, Patadia MO, Raviv J. Diagnosis of nasal airway obstruction. Otolaryngol Clin North Am 2009 Apr;42(2):207-225. [doi: 10.1016/j.otc.2009.01.004] [Medline: 19328887]
49. Nall R, Charles M. Is it strep throat? Pictures and symptoms. Medical News Today. 2017. URL: https://www.medicalnewstoday.com/articles/312433.php [accessed 2019-04-08]
50. Bhadra A, Jain M, Shidnal S. Automated detection of eye diseases. In: Proceedings of the International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET). 2016 Presented at: International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET); March 23-25, 2016; Chennai, India. [doi: 10.1109/wispnet.2016.7566355]

51. Mariakakis A, Banks MA, Phillipi L, Yu L, Taylor J, Patel SN. BiliScreen: smartphone-based scleral jaundice monitoring for liver and pancreatic disorders. Proc ACM Interact Mob Wear Ubiquit Technol 2017 Jun 30;1(2):1-26. [doi: 10.1145/3090085]
52. Wang E, Li W, Hawkins D, Gernsheimer T, Norby-Slycord C, Patel S. HemaApp: noninvasive blood screening of hemoglobin using smartphone cameras. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016 Presented at: UbiComp '16: The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12 - 16, 2016; Heidelberg Germany p. 593-604. [doi: 10.1145/2971648.2971653]
53. Seshan H, Shwetha M. Gingival inflammation assessment: image analysis. J Indian Soc Periodontol 2012;16(2):231. [doi: 10.4103/0972-124x.99267]
54. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol 1997 Sep 08;32(9):920-924. [doi: 10.3109/00365529709011203] [Medline: 9299672]
55. Shrivastava V, Londhe N, Sonawane R, Suri J. Computer-aided diagnosis of psoriasis skin images with HOS, texture and color features: a first comparative study of its kind. Comput Methods Programs Biomed 2016 Apr;126:98-109. [doi: 10.1016/j.cmpb.2015.11.013] [Medline: 26830378]
56. Kruskal W, Wallis W. Use of ranks in one-criterion variance analysis. J Am Stat Assoc 1952 Dec;47(260):583. [doi: 10.2307/2280779]
57. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Statist 1947 Mar;18(1):50-60. [doi: 10.1214/aoms/1177730491]
58. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979;6(2):65-70 [FREE Full text]
59. Nachbar F, Stolz W, Merkle T, Cognetta AB, Vogt T, Landthaler M, et al. The ABCD rule of dermatoscopy. J Am Acad Dermat 1994 Apr;30(4):551-559. [doi: 10.1016/s0190-9622(94)70061-3]
60. Trevena L, Zikmund-Fisher B, Edwards A, Gaissmaier W, Galesic M, Han P, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak 2013 Nov 29;13(2):S7. [doi: 10.1186/1472-6947-13-s2-s7]
61. Kay M, Patel S, Kientz J. How good is 85%?: a survey tool to connect classifier evaluation to acceptability of accuracy. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.: ACM Press; 2015 Presented at: CHI '15: CHI Conference on Human Factors in Computing Systems; April 18 - 23, 201; Seoul Republic of Korea p. 347-356. [doi: 10.1145/2702123.2702603]
62. McConnell JD. The price-quality relationship in an experimental setting. J Mark Res 1968 Aug;5(3):300. [doi: 10.2307/3150348]
63. Gardner DM. Is there a generalized price-quality relationship? J Mark Res 1971 May;8(2):241. [doi: 10.2307/3149770]
64. Dean DH. Brand endorsement, popularity, and event sponsorship as advertising cues affecting consumer pre-purchase attitudes. J Advert 1999 Oct;28(3):1-12. [doi: 10.1080/00913367.1999.10673585]
65. Green PE. On the design of choice experiments involving multifactor alternatives. J Consum Res 1974 Sep;1(2):61. [doi: 10.1086/208592]
66. Hammond KR, Hursch CJ, Todd FJ. Analyzing the components of clinical inference. Psychol Rev 1964;71(6):438-456. [doi: 10.1037/h0040736]
67. Cooksey RW. Judgment Analysis: Theory, Methods, and Applications. Cambridge, MA, US: Academic Press; 1996.
68. Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation. Soc Sci Med 1999 Feb;48(4):535-546. [doi: 10.1016/s0277-9536(98)00374-8]
69. Kee F, Patterson C, Wilson A, McConnell J, Wheeler S, Watson J. Judgment analysis of prioritization decisions within a dialysis program in one United Kingdom region. Med Decis Making 2016 Jun 25;22(2):140-151. [doi: 10.1177/0272989x0202200211]
70. Ryan M, Bate A, Eastmond CJ, Ludbrook A. Use of discrete choice experiments to elicit preferences. Qual Health Care 2001 Sep 01;10 Suppl 1(Supplement 1):55-60 [FREE Full text] [doi: 10.1136/qhc.0100055] [Medline: 11533440]
71. Stewart D, Shamdasani P, Rook D. Focus Groups: Theory and Practice. Thousand Oaks, CA: SAGE Publications; 2007.
72. Connor M, Sparks P. Theory of planned behaviour and health behaviour. In: Conner M, Norman P, editors. Predicting Health Behaviour: Research and Practice with Social Cognition Models. Buckingham: Open University Press; 2005:170-222.
73. Norman P, Smith L. The theory of planned behaviour and exercise: an investigation into the role of prior behaviour, behavioural intentions and attitude variability. Eur J Soc Psychol 1995 Jul;25(4):403-415. [doi: 10.1002/ejsp.2420250405]
74. Sheeran P, Orbell S. Implementation intentions and repeated behaviour: augmenting the predictive validity of the theory of planned behaviour. Eur J Soc Psychol 1999 Mar;29(2-3):349-369. [doi: 10.1002/(sici)1099-0992(199903/05)29:2/3<349::aid-ejsp931>3.0.co;2-y]
75. Bagozzi RP, Warshaw PR. Trying to consume. J Consum Res 1990 Sep;17(2):127. [doi: 10.1086/208543]
76. Sheeran P. Intention—behavior relations: a conceptual and empirical review. Eur Rev Soc Psychol 2002 Jan;12(1):1-36. [doi: 10.1080/14792772143000003]
77. Sheeran P, Webb TL. The intention-behavior gap. Soc Personal Psychol Compass 2016 Sep 05;10(9):503-518. [doi: 10.1111/spc3.12265]

78. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Appl Psychol 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]

Abbreviations

BCT: behavior change technology
CFI: comparative fit index
HBM: Health Belief Model
HCI: human–computer interaction
HIT: health intervention technology
IMC: instructional manipulation check
SRMR: standardized root mean square residual

Edited by A Kushniruk; submitted 16.05.21; peer-reviewed by M Barroso Hurtado, I Scott; comments to author 25.07.21; revised version received 15.09.21; accepted 09.10.21; published 03.01.22.

Please cite as:
Mariakakis A, Karkar R, Patel SN, Kientz JA, Fogarty J, Munson SA
Using Health Concept Surveying to Elicit Usable Evidence: Case Studies of a Novel Evaluation Methodology
JMIR Hum Factors 2022;9(1):e30474
URL: https://humanfactors.jmir.org/2022/1/e30474
doi:10.2196/30474
PMID:34982038

Original Paper

Purpose Formulation, Coalition Building, and Evidence Use in Public–Academic Partnerships: Web-Based Survey Study

Christina D Kang-Yi[1,2*], PhD; Amy Page[3*], DSW

[1]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, United States

[3]Graduate School of Social Service, Fordham University, New York City, NY, United States

[*]all authors contributed equally

Corresponding Author:
Christina D Kang-Yi, PhD
Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania
3535 Market Street, 3rd Floor
Philadelphia, PA, 19104
United States
Phone: 1 215 746 6715
Email: christina.kang-yi@pennmedicine.upenn.edu

Abstract

Background: Partnerships between academic institutions and public care agencies (public–academic partnerships [PAPs]) can promote effective policy making and care delivery. Public care agencies are often engaged in PAPs for evidence-informed policy making in health care. Previous research has reported essential partnership contextual factors and mechanisms that promote evidence-based policy making and practice in health care. However, the studies have not yet informed whether public care agency leaders' and academic researchers' perceptions of partnership purpose formulation and coalition building evolve through the PAP life cycle and whether public care agency leaders' use of research evidence differs through life cycle stages.

Objective: This exploratory study aims to focus on PAPs designed to improve youth mental health and well-being outcomes. This study also aims to identify public care agency leaders' and academic researchers' perceptions of PAP purpose formulation (structure, goals, primary function, and agenda-setting process) and coalition building (mutual benefits, trust, convener's role, member role clarity, and conflict management) by PAP life cycle stage and examine whether public care agency leaders' use of research evidence differs according to the perception of PAP purpose formulation and coalition building through the PAP life cycle.

Methods: A web-based survey of PAP experience was conducted by recruiting academic researchers (n=40) and public care agency leaders (n=26) who were engaged in PAPs for the past 10 years. Public care agency leaders additionally participated in the survey of the Structured Interview for Evidence Use scale (n=48).

Results: Most public care agency leaders and academic researchers in PAPs formed, matured, and sustained perceived their PAP as having purpose formulation context well aligned with their organizational purpose formulation context, pursuing mutual benefits, having leadership representation and role clarity, having a higher level of trust, and knowing how to handle conflicts. Most PAPs across all life cycle stages crystallized another issue to focus, but not all PAPs with issue crystallization had purpose reformulation. Public care agency leaders who trusted academic researchers in their PAP had greater use of research evidence. Public care agency leaders in PAPs that had gone through new issue crystallization also showed greater use of research evidence compared with those that had not.

Conclusions: To promote public care agency leaders' use of research evidence, focusing on developing trusting partnerships and continuously crystallizing PAP issues are important.

International Registered Report Identifier (IRRID): RR2-10.2196/14382

(JMIR Hum Factors 2022;9(1):e29288) doi:10.2196/29288

KEYWORDS

use of research evidence; public care policy; public–academic partnership; partnership purpose formulation; partnership coalition building; youth mental health and well-being

Introduction

Background

Partnerships between academic institutions and public care agencies (public–academic partnerships [PAPs]) can promote effective public policy making and care delivery. For example, local US public health departments that formally partner with academic institutions are more likely than those not engaged in partnership with academic institutes to make evidence-based policy making and implement evidence-based interventions in health care delivery [1]. Previous studies have demonstrated the important role of PAPs in training service providers [2-5], supporting the implementation of promising evidence-based practices [3-7], and conducting systems evaluation that inform policy development and program planning [2,4]. Such partnerships have effectively responded to the need for additional, more diverse, and more inclusive mental health and child welfare services [2,3,8-10].

Although previous studies have demonstrated the positive impact of PAPs on youth mental health and well-being outcomes, few empirical studies have examined whether and how PAP contexts and mechanisms evolve through the PAP life cycle and which PAP contexts and mechanisms foster public agency leaders' use of research evidence to improve youth mental health and well-being [11]. Public mental health and child welfare agencies are expected to increase the use of evidence-based care to improve mental health and well-being of vulnerable youth [12-16]. Many public care agencies partner with academic researchers to meet these expectations. Considering the multifaceted nature of public mental health and child welfare systems in the United States [17-20], it is important to develop a better understanding of the context and mechanisms that promote successful PAPs and evidence use by policy makers to improve youth outcomes.

This study has 3 aims. First, we describe a new integrated framework to understand PAP development through the PAP life cycle and potential PAP contexts and mechanisms that foster public care agency leaders' use of research evidence. Second, we summarize the literature to provide empirical support for the integrated framework, focusing on the contexts and mechanisms of PAP purpose formulation and coalition building. Third, we report our study that comprehensively explored the relationship between PAP purpose formulation and coalition building and public care agency leaders' use of research evidence by PAP life cycle stages of formed, matured, sustained, declining, and terminated.

Key PAP Process

Although research on individual components of the partnership process has revealed important information about factors that support successful partnerships, the literature has yet to bring these components together into an integrated framework [11]. Such a framework would offer a way to examine the totality of PAPs, including the contexts in which they initiate and mature, the mechanisms that propel them forward, and the outcomes that they define and achieve at various stages in relation to public care agency leaders' use of research evidence. The integrated framework by Kang-Yi [11] introduces concrete components of partnership purpose formulation and coalition building as the key contexts and mechanisms of PAPs that lead to policy makers' use of research evidence. The framework consists of three theoretical perspectives: the social partnerships perspective [21,22], the organizational life cycle perspective [23-26], and the realist evaluation perspective [27,28].

On the basis of the social partnerships and organizational life cycle perspectives [21,25], the integrated framework posits that PAPs that continuously reformulate partnership purposes and build coalitions are likely to successfully evolve through life cycle stages of being formed, matured, and sustained. Public care agency leaders in successful PAPs (being matured or sustained compared with being just formed, declining, or terminated) are more likely to use research evidence. According to the social partnerships perspective by Waddock [21], purpose formulation processes include identifying clear goals and the primary function of partnership, creating a partnership structure, and setting partnership agenda. Coalition building processes include pursuing mutual benefits for each partner, building trust among partners, solidifying the convener's role, clarifying the roles of all parties, and managing conflict [21]. The realist evaluation perspective provides a methodology for configuring contexts, mechanisms, and outcomes to examine the interplay of partnership purpose formulation, coalition building, and public care agency leaders' use of research evidence in each PAP life cycle stage and overall evolvement of PAP [27-29]. The integrated framework emphasizes continuous purpose formulation and coalition building to adjust to changing partnership environment, sustain PAP, and promote public care agency leaders' use of research evidence.

Purpose Formulation

Agenda-Setting

One key ingredient in successful PAPs is the development of a clear purpose formulation among partners. Focusing on the needs of policy makers [2,6,30-32] and having public care agency representatives who are also skilled researchers driving the agenda-setting process are important [2].

Goals

Setting clear goals for a PAP is an important aspect of achieving and measuring success. Clear PAP goals have the power to keep partners focused on working toward positive outcomes [10]. PAPs in which goals are aligned with the goals of each partnering entity can contribute to the success and longevity of those PAPs [4-6]. PAPs with clear goals can promote the use of evidence by policy makers [30].

Primary Function

PAPs can play diverse primary functions, including generating knowledge related to the development and implementation of evidence-based policy making and practices, generalizing practices to a larger population, disseminating knowledge related to the implementation of evidence-based practices, and offering technical assistance, such as professional training and program evaluation in improving service quality and outcomes [11]. Given that public care agencies and academic institutes pursue diverse missions and primary functions, alignment in primary

functions between a public care agency and an academic research institute can influence PAP sustainability and use of evidence by policy makers.

Structure

A partnership structure involves shaping governance processes, agreements around dissemination of findings, data sharing, business arrangements, ethics approvals, determining partnership mission, and general coordination among the partners [5,7,31-34]. The degree and quality of formalized structure shapes the extent of PAP success [5,32,35].

Partnership Coalition Building

Previous studies have shown that coalition building, including mutual benefits and trust, plays a critical role in successful partnerships and in promoting the use of evidence by policy makers. The key dimensions of partnership coalition building include mutual benefits, trust, convener's role, member role clarity, and conflict management.

Mutual Benefits

Successful PAPs are found to pursue mutual benefits, such as having specific agreement that ensures strategic advantages for both parties, smoother facilitation of contracts, financial incentives for the university, conducting actionable research, offering innovative ideas, improving the quality of services, offering researchers the benefit of evaluating a new theoretical model, and facilitating knowledge translation to direct practice [10,30,34-36]. Although PAPs offer a range of mutual benefits, they are not without risk. For example, time, effort, and cost of work are costs for all parties involved [36]. Risks specific to researchers include opportunity costs of spending time on projects that may not lead to publications and the potential negative impact of a changing political environment [34,36]. Risks for policy makers include spending social capital to justify engagement in the PAP, working with researchers who might not appreciate the complexity involved in the PAP work, the potential that research outcomes might not be practical, and the unknown impact the partnership may have on the organization [34,36].

Trust

Trust is another vital component of partnerships' success. Trust plays a key role in the sustainability of partnerships, leading to continued work, additional projects, and system-level changes. Trusting relationships among partners also support PAPs in weathering leadership changes, particularly when that work has become integral to the functioning of an agency, promoting more efficient and purposeful engagement of policy makers in the research process [31,36]. Trust among partners may also facilitate the use of evidence by policy makers. For example, some PAPs appoint personnel specifically to serve as relationship cultivators and to seek input into research questions to be explored by PAPs [30].

Convener's Role

PAPs need conveners to bring partners together into a partnership formation. Previous studies have documented the importance of such a role in bringing partners together in long-standing relationships within both organizations, identifying problem areas and developing initiatives in response, maintaining the necessary structure of PAP to disseminate the information generated by the partnership, and promoting the use of research evidence by policy makers [33,35,37]. Individuals who possess knowledge spanning both research and policy realms can support translating knowledge into the policy process [35].

Role Clarity

Clear delineation of roles among partners related to developing research questions and methodology as well as the eventual dissemination of the findings is important for successful PAPs [30,33]. In addition, clear communication between partners about how decisions are to be made and whether researchers can provide policy recommendations is critical [10], as these decisions can make a difference in informing policy makers and promoting the use of research evidence among policy makers. Partnerships that are slow in building comprehensive leadership teams and having members who are unsure of their roles can delay the generation of useful evidence for policy.

Conflict Management

Conflict is not unusual in the life of a partnership. Disagreement over project aims and funding [30] and other partnership processes, such as agenda-setting and contracting, can increase. Effective conflict management skills are important in building successful PAPs that lead to the use of research evidence in policy making.

This Study: Web-Based Survey of PAPs

Our study aims to focus on PAPs designed to improve youth mental health and well-being. This study also aims to identify whether contexts and mechanisms of PAP partnership purpose formulation (structure, goals, and primary function as contexts and agenda-setting process as mechanism) and coalition building (convener's role, leadership representation, role clarity, and conflict management as contexts and mutual benefits and trust as mechanisms) evolve through PAP life cycle stages (formed, matured, sustained, declining, and terminated). We also examined whether public care agency leaders' use of research evidence differs according to their perception of the PAP life cycle stage, purpose formulation, and coalition building. Research evidence was defined as relevant conceptual frameworks or reviews and empirical findings from systematic qualitative, quantitative, or mixed research methods projects [38]. The study was approved by the institutional review board of the University of Pennsylvania (see Kang-Yi [11] for the published study protocol).

Methods

Sampling and Participant Recruitment

A web-based survey of PAP partnership experience and use of research evidence was conducted by recruiting academic researchers and public care agency leaders who were engaged in PAPs. See Page et al [39] for a detailed discussion of the approach used to identify PAP researchers and public care agency leaders. To recruit public care agency leaders and academic researchers who were engaged in PAPs, we identified

PAPs through two primary methods: a web-based search of peer-reviewed journals and Google for key terms related to youth-focused PAPs and national and local meetings of professionals and researchers in the fields of mental health and child welfare. A total of 87 PAPs were identified, which met the following criteria: formed on a project, program, or intervention basis or as a consortium; aimed to improve mental health and well-being outcomes for youth aged 12-25 years; and comprised at least one or more state or local county mental health and child welfare agencies and one or more academic researchers. PAPs focused on youth outside the United States or established outside the United States and PAPs terminated within 10 years before the study initiation in 2017 were excluded. Of the 87 PAPs identified, we reached out to at least one public care agency leader in 67 PAPs and at least one academic researcher in 83 PAPs.

Once we identified researchers and public care agency leaders, we emailed them a link to the web-based survey along with introductory information about the nature of the study. A US $35 gift card was offered for full completion of a survey. Respondents were informed that the link was unique to them and asked not to share it with others. Data were collected from March 2019 to February 2020. The survey was tested for usability and accuracy by the research team and a small number of colleagues before being shared with potential respondents. In addition, the Checklist for Reporting Results of Internet E-Surveys [40] was used to report the survey as needed.

Survey Measures

To respond to the survey questionnaire, the participants were asked to focus on the latest PAP or one of the PAPs for the past decade if they were not engaged in PAP at the time of the survey. The Structured Interview for Evidence Use (SIEU) [41] was used to identify public care agency leaders' engagement level of research evidence, which refers to the frequency of using various types of sources for research evidence; public care agency leaders' ratings of the importance of evaluating the validity, reliability, and relevance of research evidence; and various factors leading public care agency leaders to use or ignore research evidence in deciding to adopt a new program or intervention. The SIEU was developed based on the posit that research use is driven by context and social relationships [41]. Thus, SIEU as a tool reflects the integrated conceptual framework being tested in this study. SIEU includes input, process, and output scales. The input scale (20 items) assesses the source of research evidence that public care agency leaders obtain. The process scale assesses how public care agency leaders evaluate the research evidence obtained and includes 3 subscales of self-assessment for validity and reliability of research evidence (10 items), reliance on others (5 items), and self-assessment for relevance (5 items). The output scale (20 items) assesses whether public care agency leaders eventually use the research evidence or ignore the evidence. The measurement responses use a 5-point Likert-type scale ranging from 1 (not at all) to 5 (all the time) for the items contained in the input scale and a 5-point Likert-type scale ranging from 1 (not important) to 5 (very important) for the items contained in the process and output scales. Each subscale measure and the total SIEU score are represented as average scores. Higher scores indicate higher agreement with the sources of evidence obtained for the input scale, more frequent evaluation of research evidence for the process scale, and greater use of research evidence for the output scale. SIEU has shown high internal consistency reliability (Cronbach α=.88) [41].

The PAP experience survey was developed for this study [11]. The questionnaire included 41 questions that were based on the potential PAP context, mechanism, and outcome configuration developed for the study [11,39]. These questions included both a Likert-type scale and open-ended questions. The survey items focused on the following four areas: (1) partnership purpose formulation (structure, goals, primary function, and agenda-setting process), (2) perceptions of partnership coalition building (mutual benefit, trust, convener's role, leadership representation, role clarity, and conflict management), (3) perception of the PAP life cycle stage, and (4) public care agency leaders' use of research evidence. We built and administered the web-based survey in the Research Electronic Data Capture [42], a secure web-based data collection tool that includes data entry forms and web surveying features.

A total of 48 public care agency leaders participated in the web-based SIEU survey scale [41], and 40 academic researchers and 26 public care agency leaders participated in the PAP experience survey. The survey response rates were 72% (48/67) for the SIEU survey, 48% (40/83) for academic researchers' PAP experience survey, and 39% (26/67) for public care agency leaders' PAP experience survey.

Analysis

The reliability of the SIEU was calculated using Cronbach α internal consistency for each of the subscales and the overall scale. Frequencies, percentages, and mean scores were calculated to identify (1) public care agency leaders' and academic researchers' ratings of alignment between PAP structure, goals, primary function, and agenda-setting process and their organizational structure, goals, primary function, and agenda-setting process by PAP life cycle stage (formed, matured, sustained, declining, and terminated); (2) public care agency leaders' and academic researchers' ratings of PAP coalition building (mutual benefits, trust, convener's role, leadership representation, role clarity, and conflict management) by PAP life cycle stage; (3) public care agency leaders' and academic researchers' ratings of their partnership outcomes (identifying another issue to focus on and reformulate PAP purpose); and (4) public care agency leaders' use of research evidence by the ratings of PAP life cycle stage, purpose formulation, and coalition building.

The original study design [11] was to recruit academic researchers and public care agency leaders in pairs. However, because of the low response rate for the PAP experience survey, we conducted group-level analysis for public care agency leaders and academic researchers, respectively, instead of conducting the analysis in pairs.

Results

Demographic Characteristics and Work Experience of Study Participants

As shown in Multimedia Appendix 1, public care agency leaders' age and years of experience in the fields were distributed evenly. Of the public care agency leaders who answered the demographics and work experience questions, most (20/31, 65%) held a master's degree. More than two-thirds of the public care agency leaders were women (21/31, 68%) and White (22/31, 71%). More than two-thirds of public care agency leaders (21/31, 68%) reported being at their current organizations for more than 10 years. More than three-fourths of public care agency leaders (24/31, 77%) had been involved in their current PAP for fewer than 10 years. The PAP roles they played were diverse and distributed evenly, and most (23/31, 74%) of public care agency leaders reported having been engaged in 5 or fewer PAPs.

As shown in Multimedia Appendix 2, academic researchers' age and years of experience were also evenly distributed as were years at the current organization. Of the academic researchers who answered the demographic and work experience questions, most were women (30/40, 75%) and White (33/40, 83%) and held a doctoral degree (26/40, 65%). More than one-fourth of the academic researchers (11/40, 28%) had been involved with their current PAP for more than 10 years. The PAP roles they played were diverse, and only under one-third (11/40, 28%) identified their role as principal investigator, lead evaluator, and university lead.

PAP Life Cycle Stages and SIEU Scale Scores

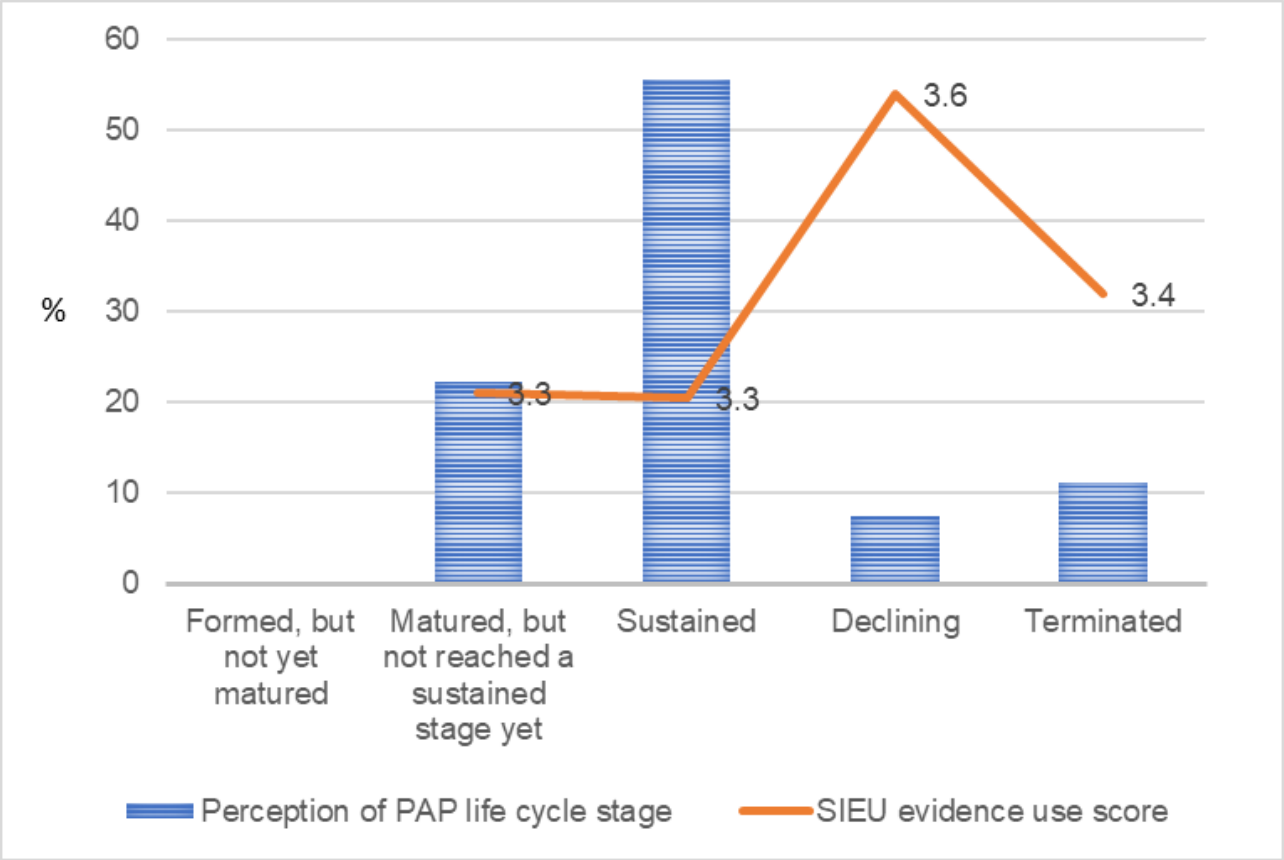
The average total SIEU score was 3.1 (SD 0.81; range 0.9-4.1). The internal consistency reliability of the SIEU based on the study sample was high (Cronbach α=.89). The mean score for the SIEU Input scale, the assessment of the source of research evidence that public care agency leaders obtain, was 2.9 (SD 0.46; range 1.8-3.9). The internal consistency reliability of the input scale was a Cronbach α value of .80. The mean SIEU process scale, the assessment of how public care agency leaders evaluate research evidence obtained, was 3.8 (SD 0.68; range 0.4-4.8). The internal consistency reliability of the process scale was a Cronbach α value of .85. The mean SIEU output scale, the assessment of public care agency leaders' use of research evidence, was 3.1 (SD 0.74; range 0.2-3.9). The internal consistency reliability of the output scale was a Cronbach α value of .74.

As shown in Figure 1, 56% (15/26) of the public care agency leaders answered that their PAP was in a sustained stage, 22% (6/26) answered that their PAP was matured but did not reach a sustained stage yet, 11% (3/26) answered that their PAP was terminated at the time of the survey, 7% (2/26) answered that their PAP was declining, and 4% (1/26) answered that they were unsure about the stage of their PAP life cycle. None of the public care agency leaders answered that their PAP was formed but not matured yet.

For academic researchers, 45% (18/40) of the academic researchers answered that their PAP was in a sustained stage, 18% (7/40) answered that their PAP was matured but did not reach a sustained stage yet, 18% (7/40) answered that their PAP was formed but not reached a matured stage yet, 10% (4/40) answered that their PAP was declining, and another 10% (4/40) answered that their PAP was terminated.

For public care agency leaders' use of research evidence (Figure 1), the public care agency leaders who answered that their PAP was declining had the highest SIEU output score (mean score 3.6, SD 0.42), followed by those who answered that their PAP was terminated (mean score 3.4, SD 0.13), those who answered that their PAP was mature but did not reach a sustained stage yet (mean score 3.3, SD 0.24), and those who answered that their PAP was sustained (mean score 3.3, SD 0.32).

Figure 1. Public care agency leaders' perception of public–academic partnership life cycle stage and the Structured Interview for Evidence Use score. PAP: public–academic partnership; SIEU: Structured Interview for Evidence Use.



Perceptions of Purpose Formulation Context and Mechanism (Primary Function, Goals, Structure, and Agenda-Setting Process) and PAP Life Cycle Stage

As shown in Table 1, for PAPs in matured and sustained stages, only one public care agency leader in each group (1/6, 17% and 1/7, 7% of the PAPs, respectively) perceived the primary function of their PAP as perfectly aligned with the primary function of their organization. None of the public care agency leaders in PAPs declining and PAPs terminated perceived perfect alignment. A total of 4 academic researchers in formed, matured, and sustained PAPs (1/7, 14%; 1/7, 14%; and 2/17, 12% of the PAPs, respectively) perceived the primary function of partnership as perfectly aligned with the primary function of their organization.

Regarding the alignment of structures between PAP and partnering organizations, more than 86% (57/66) of both public care agency leaders and academic researchers answered that the structures were fairly well to perfectly aligned across all PAP life cycle stages. Three of the academic researchers in declining and terminated PAPs (2/4, 50% and 1/4, 25% of the PAPs, respectively) perceived very little alignment in the structures, whereas all public care agency leaders in PAPs declining and terminated perceived very well or perfectly well-aligned structures. All public care agency leaders perceived their PAP goals as fairly well to perfectly aligned with their organizational goals across all PAP life cycle stages. On the other hand, 2 of the academic researchers in PAPs declining and terminated (2/8, 25% of the PAPs) perceived their PAP goals as very little aligned with their organizational goals.

As shown in Table 2, 3 of the 5 public care agency leaders in the PAPs declining and terminated, perceived their PAP agenda-setting process as not at all or very little driven by the public care agency leaders. Academic researchers' perception was similar; 3 of the 15 academic researchers in the PAPs formed, declining, and terminated perceived their PAP agenda-setting process as not at all or very little driven by public care agency leaders. More than 97% (30/31) of academic researchers in formed, matured, and sustained PAPs perceived their PAP agenda-setting process as driven by public care agency leaders. Almost half of the public care agency leaders (n=12) perceived very little of their PAP agenda process as driven by academic researchers, and this was consistent regardless of their perception of the PAP life cycle stage. The academic researchers' perceptions were similar. Regardless of the PAP life cycle stage, almost half of academic researchers (n=18) perceived their PAP agenda-setting as not at all or very little driven by the researcher.

Table 1. Public–academic partnership purpose formulation context: perception of alignment in primary function, structure, and organizational goals (public care agency leaders [N=26] and academic researchers [N=40])[a].

Parameters	Formed, n (%)		Matured, n (%)		Sustained, n (%)		Declining, n (%)		Terminated, n (%)	
	Public care agency leaders (n=0)	Academic researchers (n=7)	Public care agency leaders (n=6)	Academic researchers (n=7)	Public care agency leaders (n=15)	Academic researchers (n=17)	Public care agency leaders (n=2)	Academic researchers (n=4)	Public care agency leaders (n=3)	Academic researchers (n=4)
Primary function[b]										
Not at all	N/A[c]	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (25)
Very little	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0(0)	0 (0)	1 (25)	0 (0)	0 (0)
Fairly well	N/A	2 (29)	2 (33)	0 (0)	3 (20)	2 (12)	0 (0)	2 (50)	1 (33)	0 (0)
Quite well	N/A	3 (43)	0 (0)	3 (42)	2 (13)	4 (24)	1 (50)	0 (0)	2 (67)	1 (25)
Very well	N/A	1 (14)	3 (50)	3 (42)	7 (47)	9 (53)	1 (50)	1 (25)	0 (0)	2 (50)
Perfectly	N/A	1 (14)	1 (17)	1 (14)	1 (67)	2 (12)	0 (0)	0 (0)	0 (0)	0 (0)
Do not know or unsure	N/A	0 (0)	0 (0)	0 (0)	1 (67)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Structure alignment[d]										
Not at all	N/A	1 (14)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Very little	N/A	1 (14)	1 (17)	0 (0)	1 (67)	0 (0)	0 (0)	2 (50)	0 (0)	1 (25)
Fairly well	N/A	1 (14)	2 (33)	1 (14)	5 (33)	2 (12)	0 (0)	1 (25)	0 (0)	1 (25)
Quite well	N/A	2 (29)	0 (0)	4 (57)	3 (20)	5 (29)	1 (50)	1 (25)	1 (33)	1 (25)
Very well	N/A	2 (29)	3 (50)	1 (14)	5 (33)	8 (47)	1 (50)	0 (0)	2 (67)	1 (25)
Perfectly	N/A	0 (0)	0 (0)	0 (0)	1 (67)	2 (12)	0 (0)	0 (0)	0 (0)	0 (0)
Do not know or unsure	N/A	0 (0)	0 (0)	1 (14)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Organizational goals[e]										
Not at all	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Very little	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (25)	0 (0)	1 (25)
Fairly well	N/A	2 (29)	1 (17)	0 (0)	4 (27)	2 (12)	0 (0)	2 (50)	1 (33)	0 (0)
Quite well	N/A	1 (14)	1 (17)	3 (43)	2 (13)	3 (18)	1 (50)	0 (0)	1 (33)	1 (25)
Very well	N/A	3 (43)	4 (67)	3 (43)	8 (53)	10 (59)	1 (50)	1 (25)	1 (33)	2 (50)
Perfectly	N/A	1 (14)	0 (0)	1 (14)	1 (67)	2 (12)	0 (0)	0 (0)	0 (0)	0 (0)
Do not know or unsure	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

[a]For each cell, the within-column percentages of public care agency leaders' and academic researchers' perceptions are presented.

[b]Response missing for public care agency leaders (n=1); response missing for academic researchers (n=1).

[c]N/A: not applicable.

[d]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=1).

[e]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=1).

Table 2. Public–academic partnership (PAP) purpose formulation mechanism (agenda-setting process; public care agency leaders [N=26] and academic researchers [N=40])[a].

Parameters	Formed, n (%)		Matured, n (%)		Sustained, n (%)		Declining, n (%)		Terminated, n (%)	
	Public care agency leaders (n=0)	Academic researchers (n=7)	Public care agency leaders (n=6)	Academic researchers (n=7)	Public care agency leaders (n=15)	Academic researchers (n=17)	Public care agency leaders (n=2)	Academic researchers (n=4)	Public care agency leaders (n=3)	Academic researchers (n=4)
Perception of PAP agenda driven by researchers[b]										
Not at all	N/A[c]	0 (0)	1 (17)	1 (14)	3 (20)	2 (12)	0 (0)	0 (0)	1 (33)	1 (25)
Very little	N/A	3 (43)	1 (17)	1 (14)	4 (27)	6 (35)	1 (50)	3 (75)	1 (33)	1 (25)
Fairly well	N/A	2 (29)	2 (33)	0 (0)	2 (13)	2 (12)	0 (0)	0 (0)	0 (0)	1 (25)
Quite well	N/A	2 (29)	1 (17)	2 (29)	1 (7)	3 (24)	1 (50)	1 (25)	0 (0)	0 (0)
Very well	N/A	0 (0)	1 (17)	3 (43)	3 (20)	1 (6)	0 (0)	0 (0)	1 (33)	1 (25)
Perfectly	N/A	0 (0)	0 (0)	0 (0)	0 (0)	1 (6)	0 (0)	0 (0)	0 (0)	0 (0)
Do not know or unsure	N/A	0 (0)	0 (0)	0 (0)	0 (0)	1 (6)	0 (0)	0 (0)	0 (0)	0 (0)
Perception of PAP agenda driven by public care agency leaders[d]										
Not at all	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (25)
Very little	N/A	1 (14)	0 (0)	0 (0)	0 (0)	0 (0)	1 (50)	2 (50)	1 (33)	0 (0)
Fairly well	N/A	0 (0)	1 (17)	1 (14)	6 (40)	3 (18)	0 (0)	2 (50)	1 (33)	1 (25)
Quite well	N/A	4 (57)	3 (50)	5 (71)	3 (20)	5 (29)	1 (50)	0 (0)	1 (33)	1 (25)
Very well	N/A	2 (29)	2 (33)	1 (14)	5 (33)	7 (41)	0 (0)	0 (0)	0 (0)	1 (25)
Perfectly	N/A	0 (0)	0 (0)	0 (0)	1 (7)	2 (12)	0 (0)	0 (0)	0 (0)	0 (0)
Do not known unsure	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

[a]For each cell, within-column percentages of public care agency leaders' and academic researchers' perception are presented, respectively.

[b]Response missing for public care agency leaders, (n=2); response missing for academic researchers (n=1).

[c]N/A: not applicable.

[d]Response missing for public care agency leaders, (n=0); response missing for academic researchers (n=1).

PAP Coalition Building Context (Convener's Role, Leadership Representation, Role Clarity, and Conflict Management) and PAP Life Cycle Stage

As shown in Table 3, most public care agency leaders and academic researchers in PAPs formed, matured, and sustained had a convener who gathered people together to carry out partnership processes, such as issue crystallization, partnership coalition building, and agenda-setting. In total, 3 of the 8 academic researchers in PAPs declining and terminated (2/4, 50% and 1/4, 25% of the PAPs, respectively) and 1 public care agency leader (1/3, 33%) in PAPs terminated answered that their PAPs were missing a convener.

Public care agency leaders' perceptions of clear leadership representation and role clarity did not differ according to the PAP life cycle stage. Approximately 27% (4/15) of public care agency leaders in PAPs sustained answered that their PAP rarely or only occasionally had leadership representation and role clarity. Overall, 30 academic researchers (30/40, 75% of all academic researchers) answered that their PAP always had leadership representation and clear roles.

Most public care agency leaders and academic researchers answered that they experienced partnership conflict across all PAP life cycle stages, except for PAPs in a formed stage. Most public care agency leaders (up to 23/26, 88%; range 67%-100% across all PAP life cycle stages) answered that PAP members knew how to manage partnership conflicts. Most academic researchers in PAPs formed, matured, and sustained (up to 29/31, 94%; range 86%-100%) answered that their PAP members knew how to handle partnership conflicts. In total, 3 of the 8 academic researchers in PAPs declining and terminated answered that their PAP members knew how to handle partnership conflicts.

As shown in Table 4, public care agency leaders' trust in researchers, academic researchers' trust in public care agency leaders, and perception of pursuing mutual benefit in partnership agenda-setting did not show meaningful patterns by the PAP life cycle stage. Researchers' perception of PAPs pursuing mutual benefit in partnership agenda-setting differed according to the PAP life cycle stage. Most researchers (30/31, 97%) in PAPs formed, matured, and sustained perceived their PAP very frequently or always pursuing mutual benefit and used to pursue

mutual benefit in setting partnership agenda. PAPs sustained had the highest percentage of academic researchers (6/17, 35%) answering their PAP always pursued mutual benefits.

As shown in Table 5, across all PAP life cycle stages, most public care agency leaders and researchers answered that their PAP resulted in focusing on another issue. Academic researchers' perception of their partnership leading to focus on another issue was the highest among PAPs matured (6/7, 86%), followed by PAPs sustained (14/17, 82%), PAPs declined (3/4, 75%), PAP terminated (2/4, 50%), and PAPs formed (3/7, 43%). More than two-thirds of the researchers (4/5, 67%) in PAPs matured answered that focusing on a new issue led to reformulating the PAP agenda-setting process. The majority of public care agency leaders (10/16, 63%) answered that the new issue did not result in reformulating the PAP agenda-setting process.

Table 3. Public–academic partnership (PAP) coalition building context (convener's role, leadership representation, role clarity, and conflict management; public care agency leaders [N=26] and academic researchers [N=40])[a].

Parameters	Formed, n (%)		Matured, n (%)		Sustained, n (%)		Declining, n (%)		Terminated, n (%)	
	Public care agency leaders (n=0)	Academic researchers (n=7)	Public care agency leaders (n=6)	Academic researchers (n=7)	Public care agency leaders (n=15)	Academic researchers (n=17)	Public care agency leaders (n=2)	Academic researchers (n=4)	Public care agency leaders (n=3)	Academic researchers (n=4)
Perception of PAP having a convener who plays the role of gathering people together[b]										
Yes	N/A[c]	6 (86)	5 (83)	5 (83)	9 (60)	14 (82)	2 (100)	2 (50)	2 (67)	1 (25)
No	N/A	1 (14)	0(0)	1 (17)	3 (20)	3 (18)	0 (0)	2 (50)	1 (33)	1 (25)
Used to have	N/A	0 (0)	1 (17)	0 (0)	2 (13)	0 (0)	0 (0)	0 (0)	0 (0)	2 (50)
Do not know or unsure	N/A	0 (0)	0 (0)	0 (0)	1 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Perception of having clear leadership representation and roles[d]										
Rarely	N/A	0 (0)	0 (0)	0 (0)	2 (13)	1 (6)	0 (0)	0 (0)	0 (0)	1 (25)
Occasionally	N/A	0 (0)	0 (0)	0 (0)	2 (13)	2 (12)	0 (0)	2 (50)	0 (0)	2 (50)
Frequently	N/A	0 (0)	1 (17)	1 (14)	1 (7)	1 (6)	0 (0)	2 (50)	1 (33)	0 (0)
Very frequently	N/A	0 (0)	2 (33)	1 (14)	1 (7)	4 (26)	0 (0)	0 (0)	1 (33)	1 (25)
Always	N/A	7 (100)	3 (50)	5 (71)	8 (53)	9 (53)	2 (100)	0 (0)	1 (33)	0 (0)
Used to have	N/A	0 (0)	0 (0)	0 (0)	1 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Do not know or unsure	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Experience of PAP conflict[e]										
Yes	N/A	1 (14)	3 (50)	4 (57)	10 (67)	8 (47)	1 (50)	4 (100)	2 (67)	4 (100)
No	N/A	6 (86)	2 (33)	3 (43)	4 (27)	9 (53)	1 (50)	0 (0)	1 (33)	0 (0)
Do not know or unsure	N/A	0 (0)	1 (17)	0 (0)	1 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Perception of PAP members knowing how to handle PAP conflicts[f]										
Yes	N/A	7 (100)	4 (67)	6 (86)	14 (93)	16 (94)	2 (100)	1 (25)	3 (100)	2 (50)
No	N/A	0 (0)	2 (33)	0 (0)	0 (0)	1 (6)	0 (0)	2 (50)	0 (0)	2 (50)
Do not know or unsure	N/A	0 (0)	0 (0)	1 (14)	1 (7)	0 (0)	0 (0)	1 (25)	0 (0)	0 (0)

[a]For each cell, the within-column percentages of public care agency leaders' and academic researchers' perceptions are presented.

[b]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=2).

[c]N/A: not applicable.

[d]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=1).

[e]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=1).

[f]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=1).

Table 4. Public–academic partnership (PAP) coalition building mechanism (mutual benefit and trust in PAP agenda-setting; public care agency leaders [N=26] and academic researchers [N=40])[a].

Parameters	Formed, n (%)		Matured, n (%)		Sustained, n (%)		Declining, n (%)		Terminated, n (%)	
	Public care agency leaders (n=0)	Academic researchers (n=7)	Public care agency leaders (n=6)	Academic researchers (n=7)	Public care agency leaders (n=15)	Academic researchers (n=17)	Public care agency leaders (n=2)	Academic researchers (n=4)	Public care agency leaders (n=3)	Academic researchers (n=4)
Perception of mutual benefit in PAP agenda setting[b]										
Rarely	N/A[c]	0 (0)	1 (17)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (25)
Occasionally	N/A	1 (14)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (75)	1 (33)	1 (25)
Frequently	N/A	1 (14)	0 (0)	3 (50)	5 (33)	3 (18)	0 (0)	0 (0)	0 (0)	2 (50)
Very frequently	N/A	2 (29)	0 (0)	2 (33)	2 (13)	8 (47)	0 (0)	1 (25)	1 (33)	0 (0)
Always	N/A	3 (43)	4 (67)	1 (17)	7 (47)	6 (35)	2 (100)	0 (0)	1 (33)	0 (0)
Used to pursue mutual benefit	N/A	0 (0)	1 (17)	0 (0)	1 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Do not know or unsure	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Perception of the level of trust academic researchers have for public care agency leaders[d]										
High	N/A	5 (86)	4 (67)	5 (71)	7 (47)	14 (83)	1 (50)	1 (25)	2 (67)	1 (25)
Moderate	N/A	1 (14)	1 (17)	1 (14)	5 (33)	2 (17)	1 (50)	1 (25)	1 (33)	1 (25)
Low	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (50)	0 (0)	1 (25)
Used to have high level of trust	N/A	0 (0)	0 (0)	0 (0)	2 (13)	0 (0)	0 (0)	0 (0)	0 (0)	1 (25)
Do not know or unsure	N/A	0 (0)	1 (17)	1 (14)	1 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Perception of the level of trust public care agency leaders have for academic researchers[e]										
High	N/A	5 (71)	5 (83)	5 (63)	7 (47)	15 (83)	1 (50)	0 (0)	2 (67)	1 (25)
Moderate	N/A	2 (29)	0 (0)	1 (13)	5 (33)	2 (17)	1 (50)	3 (75)	1 (33)	2 (50)
Low	N/A	0 (0)	0 (0)	0 (0)	1 (7)	0 (0)	0 (0)	1 (25)	0 (0)	1 (25)
Used to have high level of trust	N/A	0 (0)	0 (0)	1 (13)	2 (13)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Do not know or unsure	N/A	0 (0)	1 (17)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

[a]For each cell, the within-column percentages of public care agency leaders' and academic researchers' perceptions are presented.

[b]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=2).

[c]N/A: not applicable.

[d]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=2).

[e]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=0).

Table 5. Public–academic partnership (PAP) purpose formulation and coalition building outcome (new issue to focus and reformulation of PAP agenda-setting process; public care agency leaders [N=26] and academic researchers [N=40])[a].

Parameters	Formed, n (%)		Matured, n (%)		Sustained, n (%)		Declining, n (%)		Terminated, n (%)	
	Public care agency leaders (n=0)	Academic researchers (n=7)	Public care agency leaders (n=6)	Academic researchers (n=7)	Public care agency leaders (n=15)	Academic researchers (n=17)	Public care agency leaders (n=2)	Academic researchers (n=4)	Public care agency leaders (n=3)	Academic researchers (n=4)
Perception of PAP leading to focus on another issue[b]										
Yes	N/A[c]	3 (43)	3 (50)	6 (86)	10 (67)	14 (82)	1 (50)	3 (75)	2 (67)	2 (50)
No	N/A	4 (57)	2 (33)	1 (14)	4 (27)	2 (12)	1 (50)	1 (25)	1 (33)	2 (50)
Do not know or unsure	N/A	0 (0)	1 (17)	0 (0)	1 (7)	1 (6)	0 (0)	0 (0)	0 (0)	0 (0)
Perception of PAP leading to reformulate PAP agenda-setting process[d]										
Yes	N/A	1 (33)	2 (67)	4 (67)	4 (40)	5 (36)	0 (0)	0 (0)	0 (0)	1 (50)
No	N/A	2 (67)	1 (33)	2 (33)	6 (60)	9 (64)	1 (100)	2 (67)	2 (100)	1 (50)
Do not know or unsure	N/A	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (33)	0 (0)	0 (0)

[a]For each cell, the within-column percentages of public care agency leaders' and academic researchers' perceptions are presented.

[b]Response missing for public care agency leaders (n=0); response missing for academic researchers (n=1).

[c]N/A: not applicable.

[d]Response missing for public care agency leaders (n=10); response missing for academic researchers (n=12).

PAP Purpose Formulation and Coalition Building and Public Care Agency Leaders' Use of Research Evidence

Figures 1-9 present public care agency leaders' perceptions of PAP purpose formulation and coalition building and their use of research evidence. The average SIEU output scale score that indicates public care agency leaders' actual use of research evidence was the highest among the PAPs declining followed by PAPs terminated, PAPs formed, and PAPs matured (3.6, SD 0.42; 3.4, SD 0.13; 3.3, SD 0.32; and 3.3, SD 0.24, respectively). The average SIEU output scale score was higher in PAPs, which resulted in another issue to focus compared with the score of PAPs without issue recrystallization (SIEU scores 3.4, SD 0.28 vs 3.1, SD 0.25).

On the other hand, the SIEU output scale score did not show a correlated pattern with public care agency leaders' perceptions of the agenda-setting process. Public care agency leaders who reported that their partnering researchers used to have trust in PAP leaders (public care agency leaders) showed the highest average SIEU output scale score (3.5, SD 0.21). The SIEU output scale scores did not show correlated pattern with the public care agency leaders' perception of PAP seeking mutual benefit.

Figure 2. Public care agency leaders' perception of goal alignment and the Structured Interview for Evidence Use score. SIEU: Structured Interview for Evidence Use.

Figure 3. Public care agency leaders' perception of primary function alignment and the Structured Interview for Evidence Use score. SIEU: Structured Interview for Evidence Use.

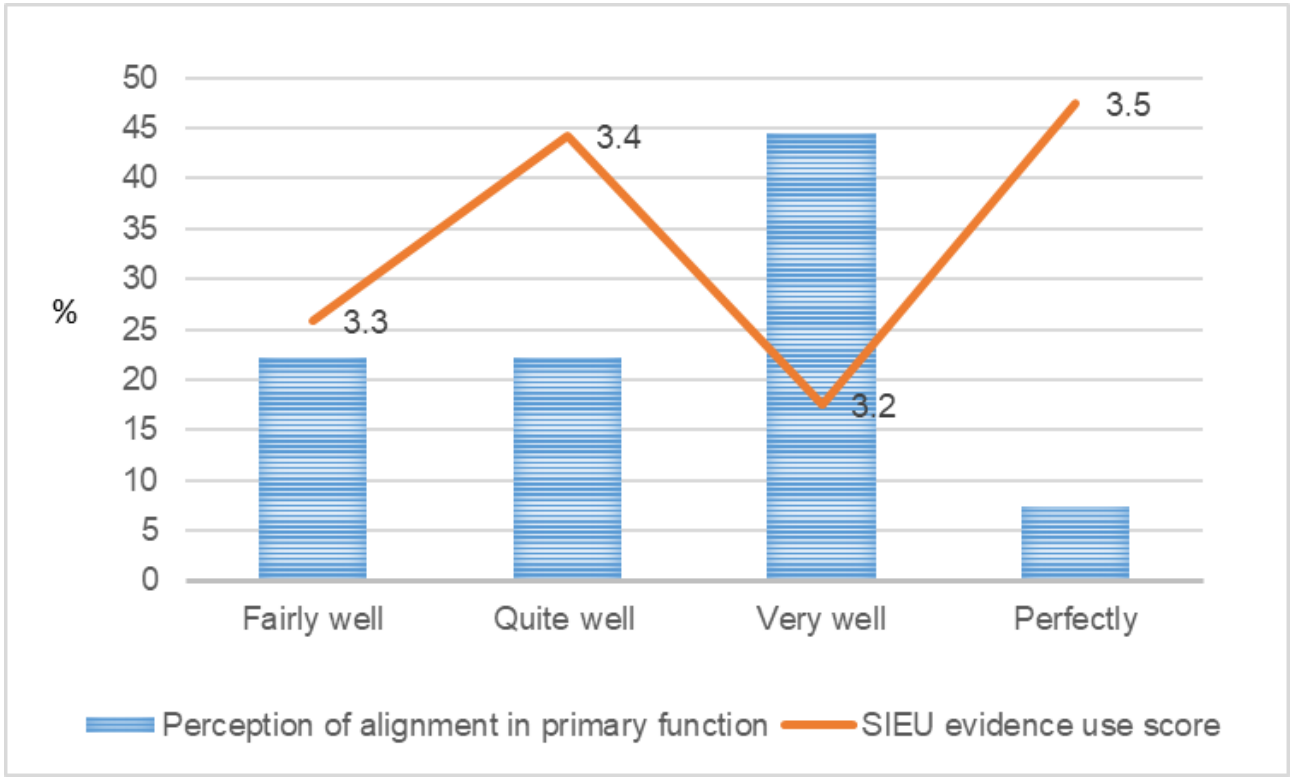


Figure 4. Public care agency leaders' perception of structure alignment and the Structured Interview for Evidence Use score. SIEU: Structured Interview for Evidence Use.

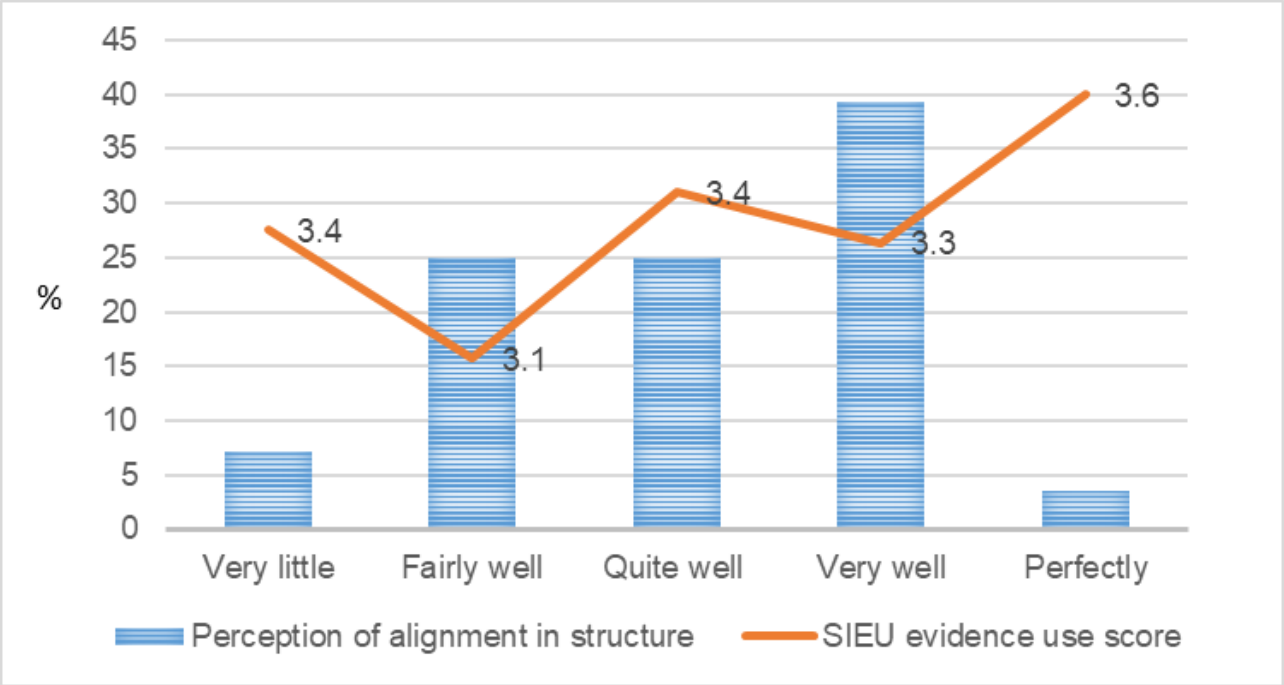


Figure 5. Public care agency leaders' perception of agenda-setting driven by public care agency leaders and the Structured Interview for Evidence Use score. SIEU: Structured Interview for Evidence Use.

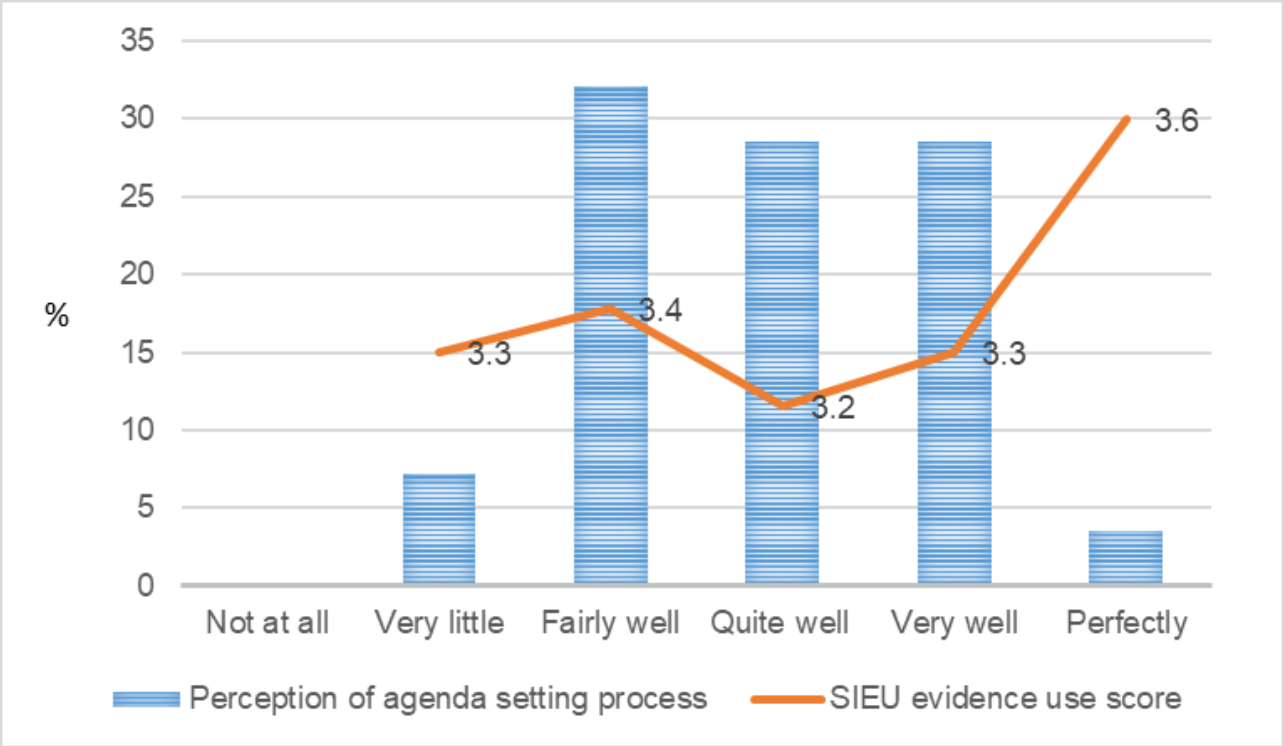


Figure 6. Public care agency leaders' perception of agenda-setting driven by researchers and the Structured Interview for Evidence Use score. SIEU: Structured Interview for Evidence Use.

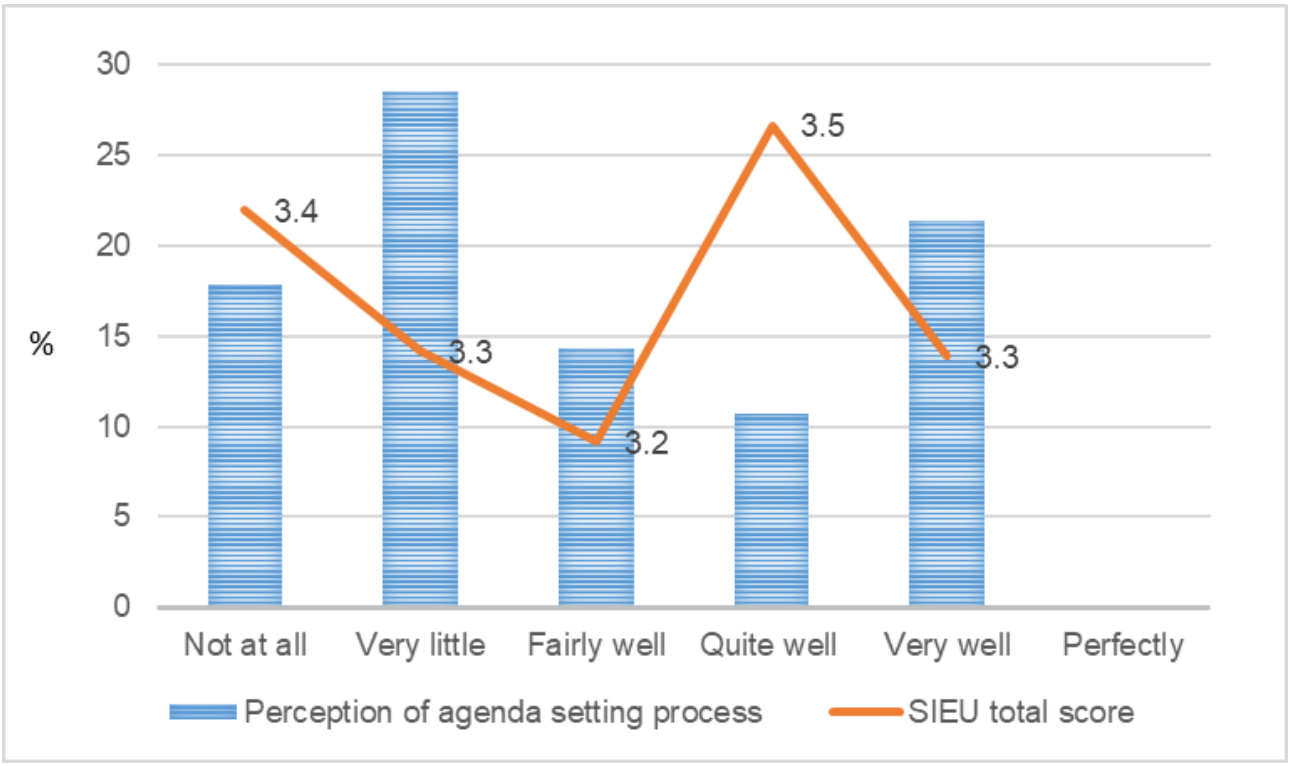


Figure 7. Public care agency leaders' response for partnership issue crystallization and the Structured Interview for Evidence Use score. PAP: public–academic partnership; SIEU: Structured Interview for Evidence Use.

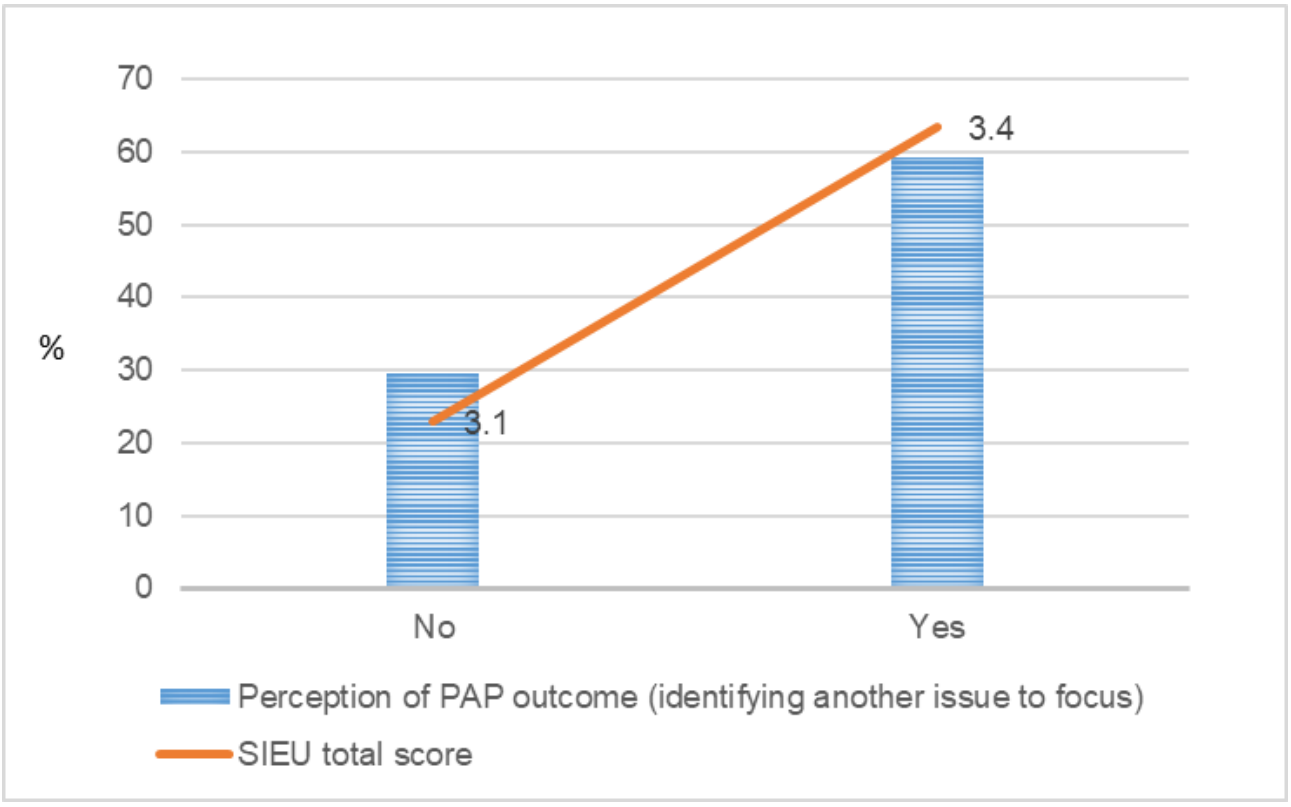


Figure 8. Public care agency leaders' response for partnership pursuing mutual benefits and the Structured Interview for Evidence Use score. PAP: public–academic partnership; SIEU: Structured Interview for Evidence Use.

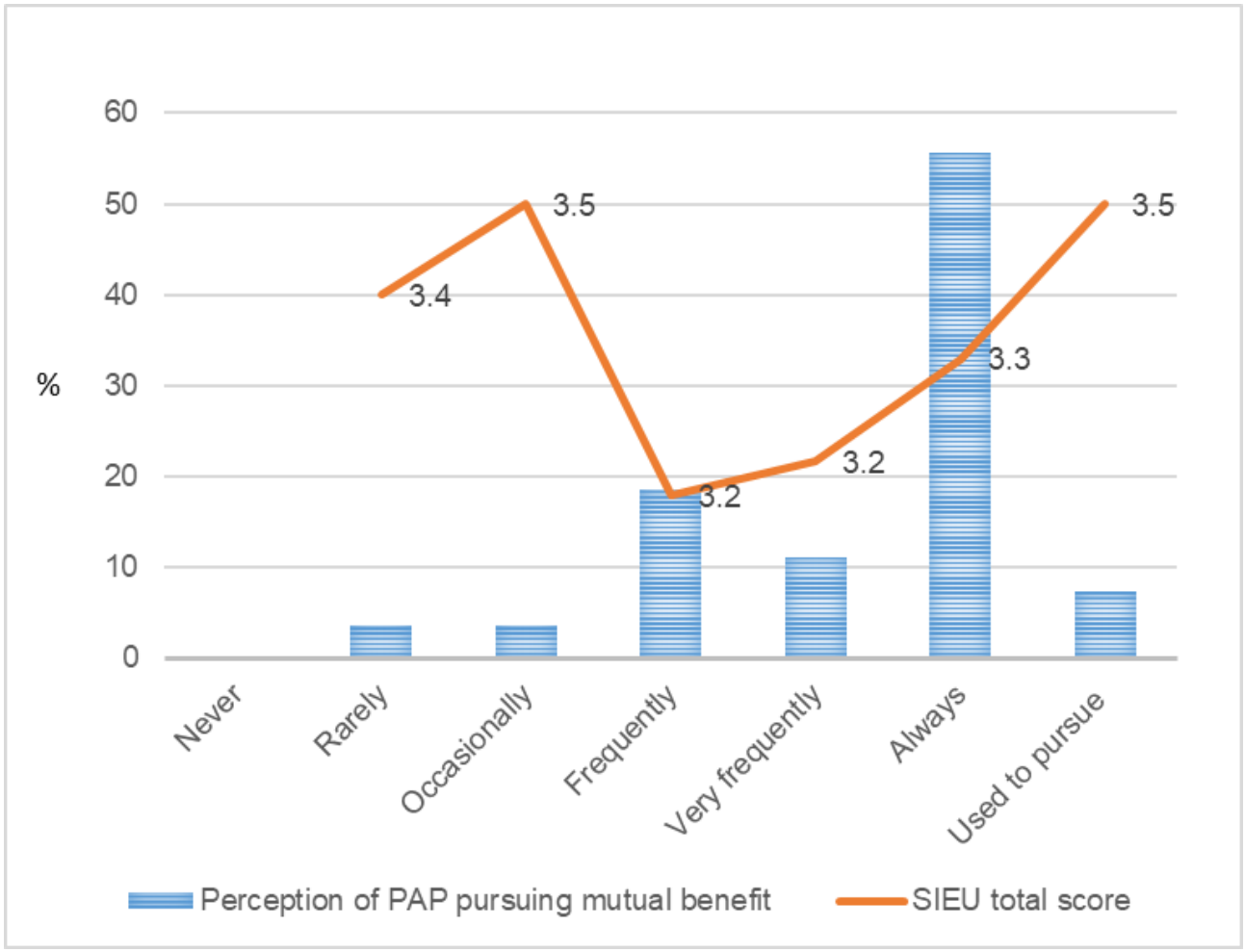
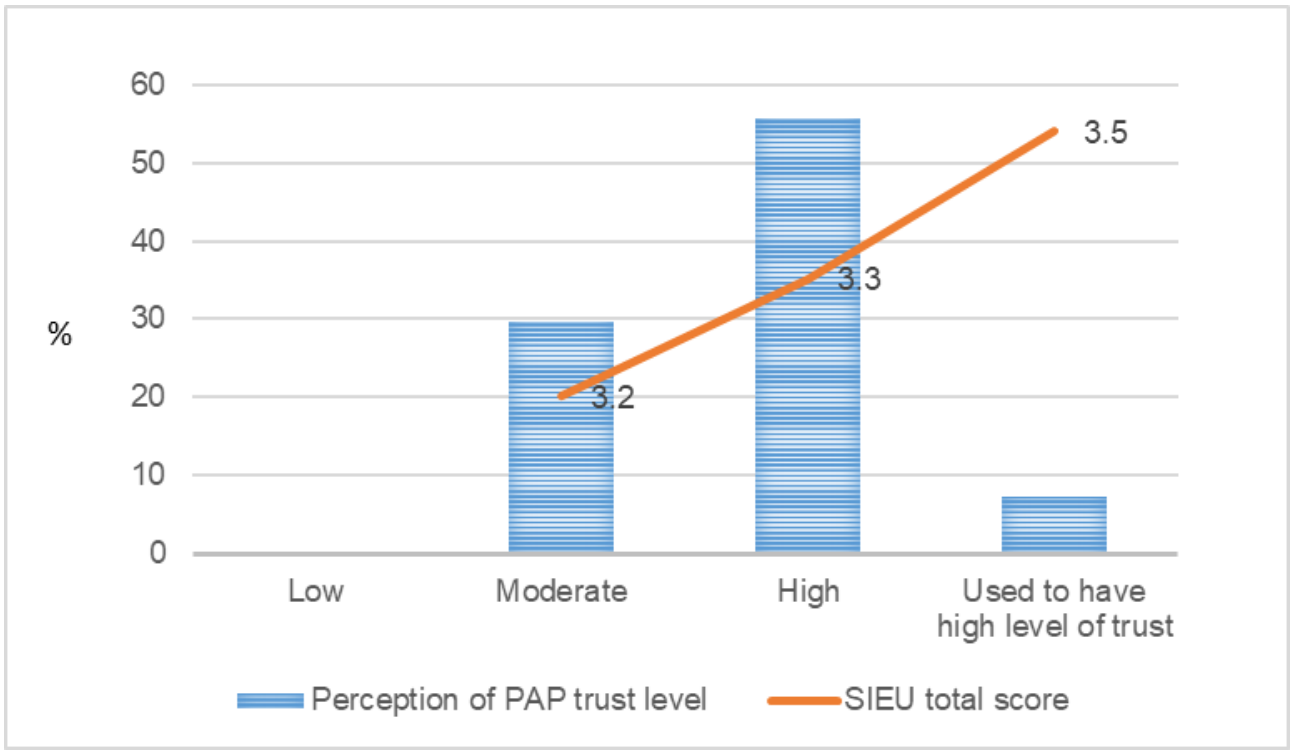


Figure 9. Public care agency leaders' perception of level of trust researchers have for the public care agency leaders and the Structured Interview for Evidence Use score. PAP: public–academic partnership; SIEU: Structured Interview for Evidence Use.



Discussion

Purpose Formulation, Coalition Building, and PAP Life Cycle Stages

The study findings show that overall, PAP purpose formulation including goals, primary function and structure, and partnership coalition building, including mutual benefits, trust, convener's role, leadership representation, role clarity, and conflict management, are important contexts and mechanisms for PAPs to evolve through life cycle stages. For the partnership contexts and mechanisms, PAPs matured were perceived more positively than PAPs formed, and PAPs sustained were perceived more positively than PAPs matured by public care agency leaders and academic researchers. However, not all the contexts and mechanisms of purpose formulation and coalition building showed evolving through the PAP life cycle stages.

Most public care agency leaders and academic researchers in PAPs formed, matured, and sustained perceived the context of partnership purpose formulation as well aligned with those of their organization. Public care agency leaders and academic researchers in PAPs declining and terminated perceived a low level of alignment in the context. This echoes the findings from studies focused on PAPs in other fields, such as public health insurance [31], environmental health [32], health care delivery [30], child welfare and mental health services [10], and general and adult mental health care [2,4-6,37] in which successful PAPs were reported to have aligned structure, goals, and agenda-setting process.

More than 40% (30/66) of public care agency leaders and academic researchers in PAPs sustained perceived that PAP agenda-setting was not at all or little driven by them. Particularly, more than one-third of researchers perceived that their PAP agenda-setting was not at all driven by academic researchers. More than 50% (25/45) of public care agency leaders and academic researchers in PAP matured and sustained perceived their PAP as always having leadership presentation and role clarity. As demonstrated in previous studies [10,34], a continuous role clarity process that responds to changing environments and needs of the mental health, child welfare, and public health fields is important for PAPs to sustain. PAPs sustained are likely to have overcome periodic leadership shifts and changes in the political environment, successfully engaging new leaders in the partnership process and continuously clarifying the roles of the members of PAP [10].

Effective conflict management skills have been shown to be important in building successful PAPs in health care delivery [30]. Most public care agency leaders and researchers experienced partnership conflict regardless of the PAP life cycle stage, except for the researchers in PAPs formed. Most public care agency leaders and academic researchers in PAPs formed, matured, and sustained reported that their PAP members knew how to handle partnership conflicts.

We found that most PAPs across all life cycle stages crystallized another issue, but the issue of crystallization did not lead to purpose reformulation for most PAPs. Although partnerships are expected to constantly review and reformulate purpose and scan their environmental changes to increase their sustainability [21], it is possible that focusing on another issue does not require changes in the PAP agenda-setting process. We did not have information on whether the new issue crystallization required PAPs to reformulate partnership purpose. Further research on specific PAP context and mechanisms that result in PAP purpose reformulation will lead to gaining an in-depth understanding.

Public Care Agency Leaders' Use of Research Evidence by Perception of PAP Purpose Formulation, Coalition Building, and PAP Life Cycle Stage

Supporting the previous research [29] on context and mechanisms for successful PAPs, our study found that developing trusting relationships with public care agency leaders and continuously crystallizing PAP issues play an important role in not only increasing PAP sustainability but also fostering public care agency leaders' use of research evidence. Public care agency leaders using research evidence may be more open to new ideas proposed by academic researchers and actively pursue issue crystallization. PAPs that continuously crystallize issues are also likely to lead public care agency leaders to be frequently exposed to research evidence. Public care agency leaders who reported their PAP as having a high level of trust in their partnering researchers also showed greater use of research evidence.

Unlike the previous research in health care delivery [30] that reports identifying clear and aligned goals as promoting partners' prioritization of their work and eventual use of evidence, our study did not find greater use of research evidence among public care agency leaders who perceived their PAP goals, primary function, and structure well aligned with their organizational goals, primary function, and structure. Previous research on health care delivery [30] and public health [36] have reported a positive relationship between PAPs seeking mutual benefit and public care agency leaders' use of research evidence. However, our study did not find the positive relationship. Public care agency leaders' use of research evidence did not show a consistent pattern by the PAP life cycle stage. Public care agency leaders who perceived their PAP as declining showed the highest level of use of research evidence. This may be attributed to the small sample size, and further research is warranted. Future research with a larger study sample and mixed methods will provide further insights.

Limitations

Our study has a few limitations. The number of public care agency leaders who participated in the PAP experience survey was limited to 26. We described in the informed consent that information provided by study participants would remain in a secure web-based database that only the key research staff could access, and that data would be analyzed at the aggregate level. Despite the statement of confidentiality and privacy protection written in the informed consent, the response rate from public care agency leaders was low. Some of the contexts and mechanisms of PAP purpose formulation and coalition building not varying by PAP life cycle stage may be attributed to the small sample size. Academic researchers' and public care agency leaders' PAP partnership experience were not analyzed in pairs because of the small sample size. Thus, our findings do not

reflect the concordance level in the perception of academic researchers and public care agency leaders in pairs. The study findings may reflect social desirability bias from the respondents. For example, as noted by Ross et al [36], researchers may have reported on positive aspects of the relationships with public care agency leaders to avoid damaging connections, and policy makers might have reported stronger reliance on evidence use because of public emphasis on evidence use. Some of the PAP contexts, such as funding opportunities and mental health and child welfare policies at the federal and state levels, are expected to influence PAP sustainability and public care agency leaders' use of research evidence. In this study, we focused on the contexts and mechanisms that can be applied to all PAPs in the fields instead of reviewing and interpreting PAP-specific contexts. A case analysis that incorporates PAP-specific contexts along with the purpose formulation and partnership coalition building can provide in-depth insights.

Conclusions

Understanding factors that promote successful PAPs and evidence use by policy makers has the potential to improve outcomes for vulnerable youth populations served by public mental health and child welfare systems in the United States. PAPs declining can revive through making changes to adapt to continuously changing environment. Our study findings suggest that continuous trust cultivation through ongoing and clear communication and continuous issue crystallization may promote public care agency leaders' use of research evidence. Academic researchers' efforts to build trust with public care agency leaders and constantly formulate issues to meet the needs of public care agency leaders who constantly experience changes in the public care environment are essential. To promote mutual benefits that link to the use of research evidence, public care agencies should establish clear research and evaluation guidelines to inform researchers of expectations when initiating and forming PAPs.

Few studies have examined PAPs in the mental health and child welfare fields despite the frequent use of PAPs. Recently, there has been increased attention to PAPs in other related fields such as health care, with rapid advancement of science such as health information technology [43]. PAPs play an important role in translating research findings into innovative policies and practices. We urge academic researchers and public care agency leaders in the fields of mental health and child welfare to pay greater attention to further understanding the partnership context and mechanisms that promote innovative evidence-based policy and practice.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Public care agency leaders' demographics and work experience.
[DOCX File , 16 KB - humanfactors_v9i1e29288_app1.docx]

Multimedia Appendix 2
Academic researchers' demographics and work experience.
[DOCX File , 16 KB - humanfactors_v9i1e29288_app2.docx]

References

1. Erwin PC, Parks RG, Mazzucca S, Allen P, Baker EA, Hu H, et al. Evidence-based public health provided through local health departments: importance of academic-practice partnerships. Am J Public Health 2019 May;109(5):739-747 [FREE Full text] [doi: 10.2105/AJPH.2019.304958] [Medline: 30896995]
2. Biegel DE, Johnsen JA, Shafran R. The Cuyahoga County Community Mental Health Research Institute: an academic-public mental health research partnership. Res Soc Work Pract 2016 Aug 18;11(3):390-403. [doi: 10.1177/1049731501011003307]
3. Beidas RS, Adams DR, Kratz HE, Jackson K, Berkowitz S, Zinny A, et al. Lessons learned while building a trauma-informed public behavioral health system in the City of Philadelphia. Eval Program Plann 2016 Dec;59:21-32 [FREE Full text] [doi: 10.1016/j.evalprogplan.2016.07.004] [Medline: 27501466]
4. Hansen EB, Seybolt DC, Sundeen SJ. Building a successful public-academic partnership to support state policy making. Psychiatr Serv 2014 Jun 01;65(6):710-712. [doi: 10.1176/appi.ps.201400067] [Medline: 24881682]
5. Stirman SW, Buchhofer R, McLaulin JB, Evans AC, Beck AT. Public-academic partnerships: the Beck Initiative: a partnership to implement cognitive therapy in a community behavioral health system. Psychiatr Serv 2009 Oct;60(10):1302-1304 [FREE Full text] [doi: 10.1176/ps.2009.60.10.1302] [Medline: 19797367]
6. Salerno A, Dixon LB, Myers RW, Smith AM, Lamberti JS, Jewell TC, et al. Public-academic partnerships: a public-academic partnership to support a state mental health authority's strategic planning and policy decisions. Psychiatr Serv 2011 Dec;62(12):1413-1415. [doi: 10.1176/appi.ps.014132011] [Medline: 22193785]
7. Truiett-Theodorson R, Tuck S, Bowie JV, Summers AC, Kelber-Kaye J. Building effective partnerships to improve birth outcomes by reducing obesity: the B'more Fit for healthy babies coalition of Baltimore. Eval Program Plann 2015 Aug;51:53-58. [doi: 10.1016/j.evalprogplan.2014.12.007] [Medline: 25547477]

8. Caplan B, Zimmet SV, Meyer EC, Friedman-Yakoobian M, Monteleone T, Jude Leung Y, et al. Prevention and recovery in early psychosis (PREP(®)): building a public-academic partnership program in Massachusetts, United States. Asian J Psychiatr 2013 Apr;6(2):171-177. [doi: 10.1016/j.ajp.2012.10.009] [Medline: 23466116]
9. Helm S, Koyanagi C, Else ', Horton M, Fukuda M. Public-academic partnerships: the University of Hawai'i Rural Health Collaboration: partnerships to provide adult telepsychiatry services. Psychiatr Serv 2010 Oct;61(10):961-963. [doi: 10.1176/ps.2010.61.10.961] [Medline: 20889630]
10. Research-practice-policy partnerships for implementation of evidence-based practices in child welfare and child mental health. William T. Grant Foundation. URL: http://wtgrantfoundation.org/library/uploads/2015/10/Research-Practice-Policy_Partnerships.pdf [accessed 2021-12-17]
11. Kang-Yi CD. Optimizing the impact of public-academic partnerships in fostering policymakers' use of research evidence: proposal to test a conceptual framework. JMIR Res Protoc 2019 May 24;8(5):e14382 [FREE Full text] [doi: 10.2196/14382] [Medline: 31127725]
12. Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003 Aug;60(8):837-844. [doi: 10.1001/archpsyc.60.8.837] [Medline: 12912767]
13. Midwest evaluation of the adult functioning of former foster youth: outcomes at age 26. Chapin Hall at the University of Chicago. 2011. URL: https://www.chapinhall.org/wp-content/uploads/Midwest-Eval-Outcomes-at-Age-26.pdf [accessed 2021-12-17]
14. Findings from the California Youth Transitions to Adulthood Study (CalYOUTH): conditions of youth at age 19. Chapin Hall at the University of Chicago. 2016. URL: https://www.chapinhall.org/wp-content/uploads/CY_YT_RE0516_4-1.pdf [accessed 2021-12-17]
15. Havlicek J, Garcia A, Smith DC. Mental health and substance use disorders among foster youth transitioning to adulthood: past research and future directions. Child Youth Serv Rev 2013 Jan;35(1):194-203 [FREE Full text] [doi: 10.1016/j.childyouth.2012.10.003] [Medline: 23766549]
16. Lewinsohn PM, Rohde P, Seeley JR. Major depressive disorder in older adolescents: prevalence, risk factors, and clinical implications. Clin Psychol Rev 1998 Nov;18(7):765-794. [doi: 10.1016/s0272-7358(98)00010-5] [Medline: 9827321]
17. How the child welfare system works. Children's Bureau. URL: https://www.childwelfare.gov/pubPDFs/cpswork.pdf [accessed 2021-12-10]
18. State vs. county administration of child welfare services. Child Welfare Information Gateway. URL: https://www.childwelfare.gov/pubPDFs/services.pdf [accessed 2021-12-10]
19. Committee for the Study of the Future of Public Health. In: The Future of Public Health. Washington (DC): National Academies Press; 1988.
20. Funding and characteristics of single state agencies for substance abuse services and state mental health agencies, 2015. Substance Abuse and Mental Health Services Administration. URL: https://store.samhsa.gov/product/Funding-and-Characteristics-of-Single-State-Agencies-for-Substance-Abuse-Services-and-State-Mental-Health-Agencies-2015/SMA17-5029 [accessed 2021-12-17]
21. Waddock SA. Understanding social partnerships: an evolutionary model of partnership organizations. Admin So 2016 Jul 26;21(1):78-100. [doi: 10.1177/009539978902100105]
22. Wymer WW, Samu S. Dimensions of business and nonprofit collaborative relationships. J Nonprofit Pub Sector Market 2003 Feb 03;11(1):3-22. [doi: 10.1300/j054v11n01_02]
23. Miller D. Evolution and revolution: a quantum view of structural change in organizations. J Manag Stud 1982 Apr;19(2):131-151 [FREE Full text] [doi: 10.1111/j.1467-6486.1982.tb00064.x]
24. Miller D, Friesen P. Archetypes of organizational transition. Administrative Sci Q 1980 Jun;25(2):268-299. [doi: 10.2307/2392455]
25. Quinn RE, Cameron K. Organizational life cycles and shifting criteria of effectiveness: some preliminary evidence. Manag Sci 1983 Jan;29(1):33-51. [doi: 10.1287/mnsc.29.1.33]
26. Whetten DA. Organizational growth and decline processes. Annu Rev Sociol 1987 Aug;13(1):335-358. [doi: 10.1146/annurev.so.13.080187.002003]
27. Pawson R, Tilley N. Realist evaluation. In: Evidence-Based Practice – Modernising the Knowledge Base of Social Work?. Germany: Verlag Barbara Budrich; 2009.
28. Pawson R. The Science of Evaluation: A Realist Manifesto. Thousand Oaks, California: SAGE Publications; 2013.
29. Jagosh J, Bush PL, Salsberg J, Macaulay AC, Greenhalgh T, Wong G, et al. A realist evaluation of community-based participatory research: partnership synergy, trust building and related ripple effects. BMC Public Health 2015 Jul 30;15:725 [FREE Full text] [doi: 10.1186/s12889-015-1949-1] [Medline: 26223523]
30. Heaton J, Day J, Britten N. Inside the "Black Box" of a knowledge translation program in applied health research. Qual Health Res 2015 Nov;25(11):1477-1491 [FREE Full text] [doi: 10.1177/1049732315580104] [Medline: 25854617]
31. Heller DJ, Hoffman C, Bindman AB. Supporting the needs of state health policy makers through university partnerships. J Health Polit Policy Law 2014 Jun;39(3):667-677. [doi: 10.1215/03616878-2682641] [Medline: 24603085]

32. Jack SM, Brooks S, Furgal CM, Dobbins M. Knowledge transfer and exchange processes for environmental health issues in Canadian Aboriginal communities. Int J Environ Res Public Health 2010 Feb;7(2):651-674 [FREE Full text] [doi: 10.3390/ijerph7020651] [Medline: 20616996]
33. Finnerty M, Neese-Todd S, Bilder S, Olfson M, Crystal S. Best Practices: MEDNET: a multistate policy maker-researcher collaboration to improve prescribing practices. Psychiatr Serv 2014 Nov 01;65(11):1297-1299 [FREE Full text] [doi: 10.1176/appi.ps.201400343] [Medline: 25756882]
34. Williamson A, Tait H, El Jardali F, Wolfenden L, Thackway S, Stewart J, et al. How are evidence generation partnerships between researchers and policy-makers enacted in practice? A qualitative interview study. Health Res Policy Syst 2019 Apr 15;17(1):41 [FREE Full text] [doi: 10.1186/s12961-019-0441-2] [Medline: 30987644]
35. Coburn AF. The role of health services research in developing state health policy. Health Aff (Millwood) 1998;17(1):139-151. [doi: 10.1377/hlthaff.17.1.139] [Medline: 9455024]
36. Ross S, Lavis J, Rodriguez C, Woodside J, Denis J. Partnership experiences: involving decision-makers in the research process. J Health Serv Res Policy 2003 Oct;8 Suppl 2:26-34. [doi: 10.1258/135581903322405144] [Medline: 14596745]
37. Brunette MF, de Nesnera A, Swain K, Riera EG, Lotz D, Bartels SJ. Public-academic partnerships: a program to improve the quality of antipsychotic prescribing in a community mental health system. Psychiatr Serv 2011 Sep;62(9):1004-1006 [FREE Full text] [doi: 10.1176/appi.ps.62.9.1004] [Medline: 21885576]
38. Tseng V. The uses of research in policy and practice and commentaries. Social Policy Report 2012 Jun;26(2):1-24. [doi: 10.1002/j.2379-3988.2012.tb00071.x]
39. Page AP, Olubiyi OB, Wong YI, Kang-Yi CD. Public-academic partnerships to foster use of research evidence in improving youth outcomes: findings from document analysis. Evid Policy 2021:1-30 (forthcoming) [FREE Full text] [doi: 10.1332/174426421x16210115966623]
40. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
41. Palinkas LA, Garcia AR, Aarons GA, Finno-Velasquez M, Holloway IW, Mackie TI, et al. Measuring use of research evidence: the structured interview for evidence use. Res Soc Work Pract 2016 Sep;26(5):550-564 [FREE Full text] [doi: 10.1177/1049731514560413] [Medline: 27616869]
42. REDCap homepage. Research Electric Data Capture. URL: https://www.project-redcap.org/ [accessed 2021-12-17]
43. Mallidou A, Dordunoo D, Borycki E, Kushniruk A, Sadeghi-Yekta K, Fraser J, et al. Perspectives and experiences of policy makers, researchers, health information technology professionals, and the public on evidence-based health policies: protocol for a qualitative study. JMIR Res Protoc 2020 Dec 17;9(12):e16268 [FREE Full text] [doi: 10.2196/16268] [Medline: 33331825]

Abbreviations

PAP: public–academic partnership
SIEU: Structured Interview for Evidence Use

Edited by A Kushniruk; submitted 31.03.21; peer-reviewed by A Mallidou, D Dordunoo; comments to author 04.07.21; revised version received 01.10.21; accepted 06.11.21; published 05.01.22.

Please cite as:
Kang-Yi CD, Page A
Purpose Formulation, Coalition Building, and Evidence Use in Public–Academic Partnerships: Web-Based Survey Study
JMIR Hum Factors 2022;9(1):e29288
URL: https://humanfactors.jmir.org/2022/1/e29288
doi:10.2196/29288
PMID:34989678

Original Paper

Iterative Development and Applicability of a Tablet-Based e-Coach for Older Adults in Rehabilitation Units to Improve Nutrition and Physical Activity: Usability Study

Lisa Happe[1], MSc; Marie Sgraja[1], BSc; Andreas Hein[1], Prof Dr Ing; Rebecca Diekmann[1], Dr oec troph

Department of Health Services Research, Assistance Systems and Medical Device Technology, Carl von Ossietzky University Oldenburg, Oldenburg, Germany

Corresponding Author:
Lisa Happe, MSc
Department of Health Services Research, Assistance Systems and Medical Device Technology
Carl von Ossietzky University Oldenburg
Haarentor Campus, Building V04
Ammerländer Heerstraße 140
Oldenburg, 26129
Germany
Phone: 49 441798 ext 2354
Email: lisa.happe@uol.de

Abstract

Background: Maintaining nutrition and exercise strategies after rehabilitation can be difficult for older people with malnutrition or limited mobility. A technical assistance system such as an e-coach could help to positively influence changes in dietary and exercise behavior and contribute to a sustainable improvement in one's nutrition and mobility status. Most apps do not provide a combination of nutrition and exercise content. In most cases, these apps were evaluated with healthy individuals aged <70 years, making transferability to vulnerable patients, with functional limitations and an assumed lower affinity for technology, in geriatric rehabilitation unlikely.

Objective: This study aims to identify the potential for optimization and enhance usability through iterative test phases to develop a nutrition and mobility e-coach suitable for older adults (≥65 years) based on individual health behavior change stages in a rehabilitation setting.

Methods: Iterative testing was performed with patients aged ≥65 years in a rehabilitation center. During testing, participants used an e-coach prototype with educational elements and active input options on nutrition and mobility as a 1-time application test. The participants performed navigation and comprehension tasks and subsequently provided feedback on the design aspects. Hints were provided by the study team when required, documented, and used for improvements. After testing, the participants were asked to rate the usability of the prototype using the System Usability Scale (SUS).

Results: In all, 3 iterative test phases (T1-T3) were conducted with 49 participants (24/49, 49% female; mean 77.8, SD 6.2 years). Improvements were made after each test phase, such as adding explanatory notes on overview screens or using consistent chart types. The use of the user-centered design in this specific target group facilitated an increase in the average SUS score from 69.3 (SD 16.3; median 65) at T1 to 78.1 (SD 11.8; median 82.5) at T3. Fewer hints were required for navigation tasks (T1: 14.1%; T2: 26.5%; T3: 17.2%) than for comprehension questions (T1: 30.5%; T2: 21.6%; T3: 20%). However, the proportion of unsolved tasks, calculated across all participants in all tasks, was higher for navigation tasks (T1: 0%, T2: 15.2%, T3: 4.3%) than for comprehension tasks (T1: 1.9%, T2: 0%, T3: 2.5%).

Conclusions: The extensive addition of explanatory sentences and terms, instead of shorter keywords, to make it easier for users to navigate and comprehend the content was a major adjustment. Thus, good usability (SUS: 80th-84th percentile) was achieved using iterative optimizations within the user-centered design. Long-term usability and any possible effects on nutritional and physical activity behavior need to be evaluated in an additional study in which patients should be able to use the e-coach with increasing independence, thereby helping them to gain access to content that could support their long-term behavior change.

(JMIR Hum Factors 2022;9(1):e31823) doi:10.2196/31823

KEYWORDS

older adults; rehabilitation; physical activity; nutrition; e-coach; usability testing; tablet computers; health behavior; mobile phone

Introduction

Background

Different demographic, clinical, biological, and lifestyle factors contribute to the development of frailty and sarcopenia in older populations. The accumulation of these risk factors leads to a reduction in resistance to health stressors. In addition to a decline in independence, there is an increased risk of falls and mortality [1,2]. In many older people, malnutrition and reduced physical activity are associated with each other [3]. The performance of physical exercises in combination with a protein-rich diet is a promising approach for the prevention and treatment of frailty [1]. Such treatments can be provided as part of outpatient or inpatient geriatric rehabilitation, although inpatient geriatric rehabilitation is more common in Germany. In the context of geriatric rehabilitation, patients are treated by a multidisciplinary team, which focuses on the individual needs and abilities of the patient [4,5]. However, long-term maintenance by older adults after rehabilitation is often unsuccessful. Negative influences on adherence, such as sudden changes in health status, a lack of interest or motivation, low self-efficacy, or low expectations of improvement [6,7] are likely to be factors affecting this lack of success.

Technical assistance systems, such as health apps, could help ensure that dietary and exercise behaviors are implemented and changed after rehabilitation. In a survey in Germany with older adults (aged >65 years) in 2020, up to 22% reported that they used health apps to track fitness data, and 16% used apps to obtain information about health, fitness, and nutrition topics. However, the proportion of seniors interviewed who could envision using such apps was more than twice as high for both types of health apps [8]. Factors that influence the acceptance of health apps should be considered in an effort to encourage older adults to use such apps. Barriers to the acceptance of health apps include a lack of trust in health apps, privacy concerns, and fear of misdiagnosis. In addition, older people who generally use apps, but who have no experience using health apps, reported a lack of health app usability and low self-confidence as reasons for poor acceptance [9]. In addition to improving older people's access to technology, it is also essential that health apps are valid, reliable, and based on current scientific evidence. It is recommended to increase the involvement of older people in the design, conceptualization, and testing of such apps [10]. Most scientifically developed health apps were evaluated with individuals aged up to 70 years or with older adults without health impairments [11-13].

Apps are one type of technical tool for improving, assisting, or supporting people. The various realizations, such as tele-visits, exergames, or health websites, can be summarized under the term eHealth. A review of the use of eHealth in the context of geriatric rehabilitation revealed that most studies (68%) involved people with neurological diseases. In addition, only 8% of all identified studies assessed the use of health apps as an eHealth intervention. However, the results on the applicability of eHealth with the target group indicated that interventions are feasible if adequate training takes place, and if the eHealth intervention is simple and has good usability [14]. For example, a study on telerehabilitation via a website compared with conventional rehabilitation after a hip fracture showed that patients in the telerehabilitation group achieved better functional scores than those in the control group [15]. In addition, in a study by Bean et al [16], the implementation of a 12-month web-based training program using a tablet and supervised by a physiotherapist led to a significant reduction in emergency department visits and hospital admissions in older adults with mobility impairments. However, there is also evidence that the more severe the health limitations of older people, the less willing they are to use eHealth [17]. A systematic review of the use of health apps to improve dietary behavior and nutrition-associated outcomes was unable to find any studies that focused on the use of such apps by people aged >70 years [18].

As previously described, a nutrition intervention in combination with an exercise intervention could lead to more significant effects than would an exercise intervention alone in older people with frailty or sarcopenia [2]. An app that offers coaching in the sense of providing information and feedback on physical activity and nutrition topics, thus supporting behavior change or the maintenance of newly adopted behaviors and tailored to the needs of older people, could therefore be a promising approach to making a rehabilitation program even more effective and sustainable. The use of the app should be continued after rehabilitation and thus provide support in everyday life in the uptake, implementation, and maintenance of the recommendations in the area of nutrition and physical activity.

To develop an age-adapted device and health app (e-coach) for older adults with deficits in nutritional status and physical activity needs, we followed the German International Organization for Standardization 9241-210:2019 *Ergonomics of human–system interaction—Part 210: Human-centered design of interactive systems* [19] and the user-centered design process [20].

These concepts first require an analysis of the context of use and, as a next step, a specification of the use requirements. The context of use was described in a previous study by performing and analyzing focus groups with older adults as well as experts [21]. In all, 3 focus groups with patients and relatives (10/17, 65% female; 16/17, 94% in a 70- to 99-year–age category) and 1 focus group with experts (2 dieticians and 1 physiotherapist) were held in a geriatric rehabilitation center. Interviews held with the focus groups were recorded, transcribed verbatim, and analyzed using content analysis. Both patients and therapists mentioned very similar points as relevant topics for e-coaches. Examples of the aspects mentioned included information about nutrition in advanced age, macronutrients, fluid intake, nutrition myths, physical activity recommendations for older adults, guidance in performing physical exercises, information on goal setting, the risk of falling, and adherence to physical activities.

However, individual perceptions of the need for further information varied widely.

The information gained was used to derive the user and design requirements for the e-coach. For geriatric patients in rehabilitation, educative content from the areas of nutrition and physical activity focused on the changes and demands of aging that should be included in the e-coach. The older adults would also like the e-coach to be able to provide them with exercises and thus support them in their training. The feedback and evaluation of input regarding nutrition and exercise are described as helpful but should not be an admonition. The results indicate that, as many patients in this age group have little experience with technology and usually use other sources of information, it is important to develop a nutrition and mobility e-coach, particularly given the easy handling and provision of clear information to individual users on the advantages of the e-coach. It is also important for older adults to avoid barriers, such as small font, low video volume, or poor contrast.

The e-coach needs to be integrated into users' daily lives without stressing or restricting them. Moreover, it must be possible to adapt the content to the physical abilities of the users, and because of the heterogeneity of older people in terms of previous knowledge and willingness to change their behavior, appropriate strategies should be used. A recent umbrella review of eHealth interventions suggested that applications involving behavior change techniques may have promising effects on physical activity, sedentary behavior, and healthy eating. However, it is not yet known which theoretical construct is the most effective [22]. In this study, the transtheoretical model of behavior change (TTM) was used as the underlying psychological construct for the e-coach. The patient was categorized into one of the five TTM phases reflecting their readiness for change: (1) precontemplation, (2) contemplation, (3) planning, (4) action, and (5) maintenance. In the first phase (precontemplation), there is no awareness of the problem or intention to act yet, whereas in the fourth stage (action), the desired behavior is already specifically being executed. Different strategies are used at each phase to achieve or sustain targeted behavior [23]. The model was first developed in the context of substance abuse treatment; however, it has also been applied to other health behavior change processes, such as increasing physical activity [24,25]. A recent systematic review on the use of the TTM in programs designed to improve physical activity in older people showed positive effects on relevant parameters, such as the reduction of sedentary behavior, the increase in activity time per week, an increase in the number of steps, and an increase in the daily total of moderate to vigorous activity time [26]. In another study that used an app to increase physical activity in healthy older adults, the TTM was also used and positive effects on exercise adherence and walking speed were observed [13].

On the basis of the findings from the focus groups, it was possible to further differentiate the settings for teaching the use of the e-coach. To give older adults time to familiarize themselves with the system and generally introduce them to its use, this introduction should already take place in the rehabilitation center. In a real health care situation, it would be easier to explain the technology to the patient; in case of questions or if further explanations are necessary, it would be more uncomplicated to address these points in a personal appointment. In addition, patients would also be able to repeat relevant content that they may not have been able to remember completely from the seminars at their own pace. Therapists should have the possibility to adapt the e-coach to the needs of the patients and their TTM phase. As, in the context of rehabilitation, the therapies take place directly between the physiotherapist or nutritionist and the patient, and the interventions are also strongly influenced by the interactions between the professionals and patients, complete automation of the e-coach would not be efficient. Adaptations of the e-coach to the patient's previous knowledge and support needs should therefore be made by a physiotherapist or a nutritionist.

Objectives

This paper aims to describe the design process and an iterative evaluation of the developed content. The aim of this study is to identify optimization potentials and enhance usability through iterative test phases to develop a nutrition and mobility e-coach based on individual health behavior change stages, usable for older adults (≥65 years) in a rehabilitation setting.

Methods

Study Design

The e-coach prototypes were evaluated with older adults in 3 iterative test phases, using a between-subject design. User experience was reflected by the System Usability Scale (SUS) [27] and participants' comments, which were made while thinking aloud during the tests.

To detect and analyze usability problems in more detail, at least 10 patients were included in each iterative test phase [28]. Then, based on the feedback, improvements were made and the prototypes were evaluated again with the target group. The opinions of other relevant stakeholders (physiotherapists and nutritionists) were taken into account throughout the design process by involving professionals from these disciplines in the study team.

Ethics Approval

The study was approved by the Ethical Review Board of the Carl von Ossietzky University Oldenburg (registration number: 2018–132). We conducted the study in accordance with the Declaration of Helsinki and the underlying data protection regulation.

Participants

Patients in rehabilitation, from geriatric and cardiology wards, were eligible based on the following inclusion criteria: (1) participants aged ≥65 years and (2) participants were able to speak and understand German. Exclusion criteria were (1) severe visual impairment (eg, inability to read large font on a screen), (2) severe hearing impairment (eg, deafness), or (3) inability to understand study information and provide informed written consent (eg, aphasia or severe cognitive impairment or dementia). Participants were recruited by placing flyers in the patients' wards.

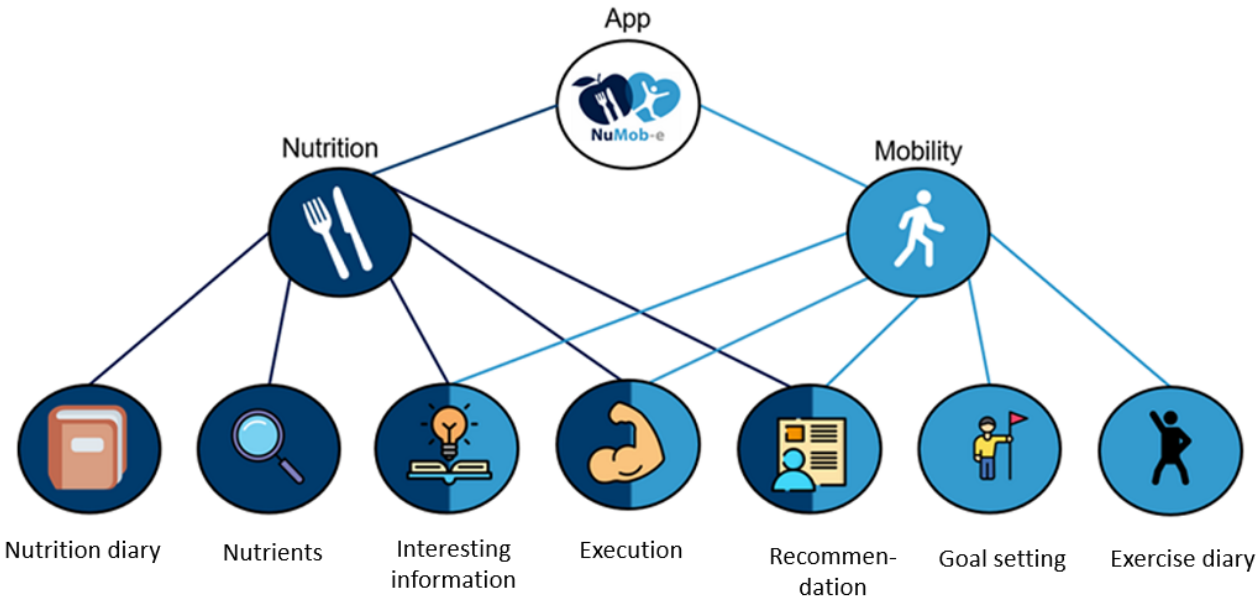
Tablet App (e-Coach)

The e-coach screens were designed in Adobe XD (version 34; Adobe Systems) for a 10-inch tablet in landscape mode. Design guidelines for apps for older adults were used to take into account specific requirements, such as a decline in vision or decreased motor abilities when developing the prototype [29,30]. To ensure a linear navigation structure, the position and design of the navigation buttons were always identical. The only gesture needed was to tap a button to minimize the number of necessary gestures. The structure of different screens containing the same type of information provision (eg, videos or texts) was kept identical.

The e-coach contained two main topics: *mobility* and *nutrition*. Both topics offered five modules in total, which were identified from previous focus group discussions with patients and experts. The e-coach was designed to support behavior change in older, vulnerable patients in rehabilitation by providing information, education about risk factors, and strategies for implementing and maintaining nutrition and physical activity recommendations. Different elements, such as videos, texts, or quizzes, were used to provide the information. In addition, the e-coach included educational elements and active input options, such as feedback on nutrient intake through entries in a nutrition diary or instructions and documentation of physical exercises, and an overview of the achievement of exercise goals. The modules entitled *interesting facts*, *recommendations*, and *execution* were elements of both main themes, although with different subthemes (eg, content about fall risk factors or strategies to promote the intake of fluids). The modules and their content differed according to the phase of the TTM. For example, by focusing more on educational content in the precontemplation and contemplation phases, strategies to increase problem awareness, environmental re-evaluation, and emotional experience are promoted, whereas, in the planning, action, and maintenance phases, more content was introduced that enabled the maintenance of new behaviors, such as enhancing self-efficacy or offering specific action guidance.

The module content and the composition of modules for each TTM phase were compiled in advance by members of the study team who have expertise in physiotherapy and nutritional therapy. The content of the nutrition modules was based on the recommendations of the European Society for Clinical Nutrition and Metabolism guidelines on clinical nutrition and hydration in geriatrics [31], as well as on the information provided in brochures issued by the IN FORM initiative of the *Deutsche Gesellschaft für Ernährung* (German Nutrition Society) entitled *Essen und Trinken im Alter* (Eating and drinking in old age) [32] and *Mangelernährung im Alter* (Malnutrition in old age) [33]. The content of the physical activity modules was mainly based on the national recommendations for physical activity and physical activity promotion, the *Älter werden in Balance* (Getting Older in Balance) program of the Federal Center for Health Education and the recommendations from the IN FORM initiative run by *Bundesarbeitsgemeinschaft der Seniorenorganisationen* (German National Association of Senior Citizens' Organizations) [34]. Physical exercises were based on exercises from the Otago Exercise Program [35], Lifestyle-integrated Functional Exercise-Program [36], and home-based older people's exercise [37] programs, which were developed especially for older people and people at risk of falling. Automated adaptation of module content was not part of the e-coach. The results of the focus groups did not indicate that this would be a requirement for therapists for such apps. Moreover, the app should be able to be used as support and an addition within the scope of therapies in rehabilitation and in further outpatient care. Therefore, the use should be embedded in the context of the therapy situation, and the assessment of the TTM phase should not be done completely automatically but by the therapist in interaction with the patient. Therefore, the TTM phase is determined as described below and then set in the app. In the context of this study, no direct association between the TTM phase and the ability to perform tasks in the app was assumed, as it was a 1-time application test with specific task instructions to the participants. The general structure of the e-coach is shown in Figure 1.

Figure 1. Structure of the e-coach modules.



Procedure

Overview

Data were collected in the form of in-person testing in patients' rooms at the rehabilitation center. After the patients had been informed of the content and the procedure of the study by a study team member (LH or MS) and had signed the informed consent, a survey was conducted on sociodemographic data, data on nutrition and physical activity, the phases of behavioral change, and technology commitment. A usability test was subsequently performed. The entire process of data collection took approximately 45 minutes to 60 minutes per patient.

Evaluation of Nutritional Data and Physical Activity

Nutritional status was assessed using the Mini Nutritional Assessment-Short Form (MNA-SF) [38] and a survey of nutritional behavior based on an eating protocol for a typical day. The data from the eating protocol were compared with the recommendations of the German Nutrition Society for people aged >65 years [39]. Physical activity was evaluated using the Physical Activity Scale for the Elderly (PASE) [40]. The intensity, type, and duration of the activities described were compared with the German national recommendations for physical activity in older adults [41].

Phases of Behavior Change (TTM)

Patients were classified separately into the TTM for physical activity and nutrition based on the data from the dietary protocol and the PASE. Patients who did not achieve the defined target criteria in the areas of physical activity or nutrition were asked whether they had thought about changing their behavior (contemplation) and, if so, whether they had already planned to do so in any specific way (planning). Depending on their answers, patients were then categorized into the phases of precontemplation, contemplation, or planning. Patients who were already performing the target behavior were asked whether they had been doing so for a short time (action) or for a longer period (maintenance). On the basis of their answers, the patients were classified into the phases of action or maintenance.

Technology Commitment

In addition, the patients' technology commitment was assessed using a questionnaire developed by Neyer et al [42]. In this questionnaire, technology commitment was measured using a 5-point Likert scale with 12 items. The items covered statements about personal contact, interest, and the use of technologies in general [42].

Usability Task

The test procedure was explained in detail, and the contents and navigation options were shown in advance. Before testing the usability task, the patients were told that the aim of the study was not to test their abilities, but the quality and usability of the e-coach [43]. Moreover, the patients were able to choose whether they wanted to be interviewed about content from one main topic only (nutrition or mobility) or about both topics. The status of development and the general structure of the mock-ups from the 2 areas were identical, but the content differed on account of topics. All elements that could be used by patients in the final e-coach were tested; however, not every screen was tested, instead, a transferability of findings was assumed.

Testing also took place in the patient's room at the rehabilitation center. During the test, the examiner (LH or MS) and the participant sat at a table. The tablet with the app could be placed on the integrated stand of the device, placed on the table, held in the hand, or laid down by the participant. The examiner read the tasks to the participant and then observed the participant.

Usability tasks in three different domains were defined and used for each iterative testing period: navigation, comprehension, and design. Navigation tasks were used to determine whether users were able to find their way through the e-coach and use the buttons correctly. In the case of quizzes, we tested whether the screens were structured such that they could be used successfully by the participants. On quiz screens, the question was highlighted at the top of the screen (eg, What is the recommended minimum number of small portions of dairy products to eat per day?) and below it, two to three answer options were shown (eg, You should eat at least two servings of dairy products per day and You should eat at least four servings of dairy products a day) along with a prompt (press the correct answer). Comprehension tasks were intended to test whether screen content and information were correctly interpreted and understood. The aim of the design questions was to identify visual barriers such as a font that was too small or an acoustic problem, such as an extremely fast rate of speech in videos. An example of questions and tasks is shown in Table 1 and Multimedia Appendices 1 and 2.

Table 1. Task types in the different iterative testing phases as the total number of tasks and percentages per iterative testing phase.

Task type	Navigation task, T1[a], n (%)	Navigation task, T2[b], n (%)	Navigation task, T3[c], n (%)	Comprehension question, T1, n (%)	Comprehension question, T2, n (%)	Comprehension question, T3, n (%)
Navigation				N/A[d]	N/A	N/A
Next screen (1 screen)	16 (57)	5 (24)	3 (30)			
Further screen (≥2 screens)	4 (14)	5 (24)	4 (40)			
Use of back button	5 (18)	5 (24)	2 (20)			
Use of different tabs (text elements)	2 (7)	2 (10)	0 (0)			
Use of the help button	0 (0)	3 (14)	1 (10)			
Use of quizzes	1 (4)	0 (0)	0 (0)			
Use of the exercise diary	0 (0)	1 (5)	0 (0)			
Comprehension	N/A	N/A	N/A			
Purpose of the screen				5 (28)	2 (20)	1 (25)
Foresight of content				3 (17)	0 (0)	1 (25)
Nutrition diary				1 (6)	0 (0)	0 (0)
Interpretation of content				7 (39)	8 (80)	2 (50)
Understanding of quizzes				2 (11)	0 (0)	0 (0)

[a]T1: iterative phase 1.

[b]T2: iterative phase 2.

[c]T3: iterative phase 3.

[d]N/A: not applicable.

Evaluation of Usability

The results from the usability tasks were reported in three different categories (success rate, number of hints, and content of hints) to evaluate usability problems in more detail. The performance of the particular task was evaluated in the categories of *successful* and *unsuccessful*. Furthermore, the content of the hints given and the number of hints given were recorded.

For the usability test, patients were instructed to simultaneously speak their thoughts aloud while performing the tasks. The concurrent think-aloud method was intended to immediately identify and specify problems for older adults using the e-coach [44]. Attention was paid not to interrupt patients while they were still thinking or looking for an answer. However, if the participant said that they were stuck or if they were obviously having difficulties, (eg, the participant looked around for help for a longer period or became increasingly nervous) a hint was given. These hints mostly consisted of a slight rephrasing of the question or a request to the participant to read through the contents of the screen again. If a hint had to be given, this was noted for the relevant task, and the number of times hints were given was counted.

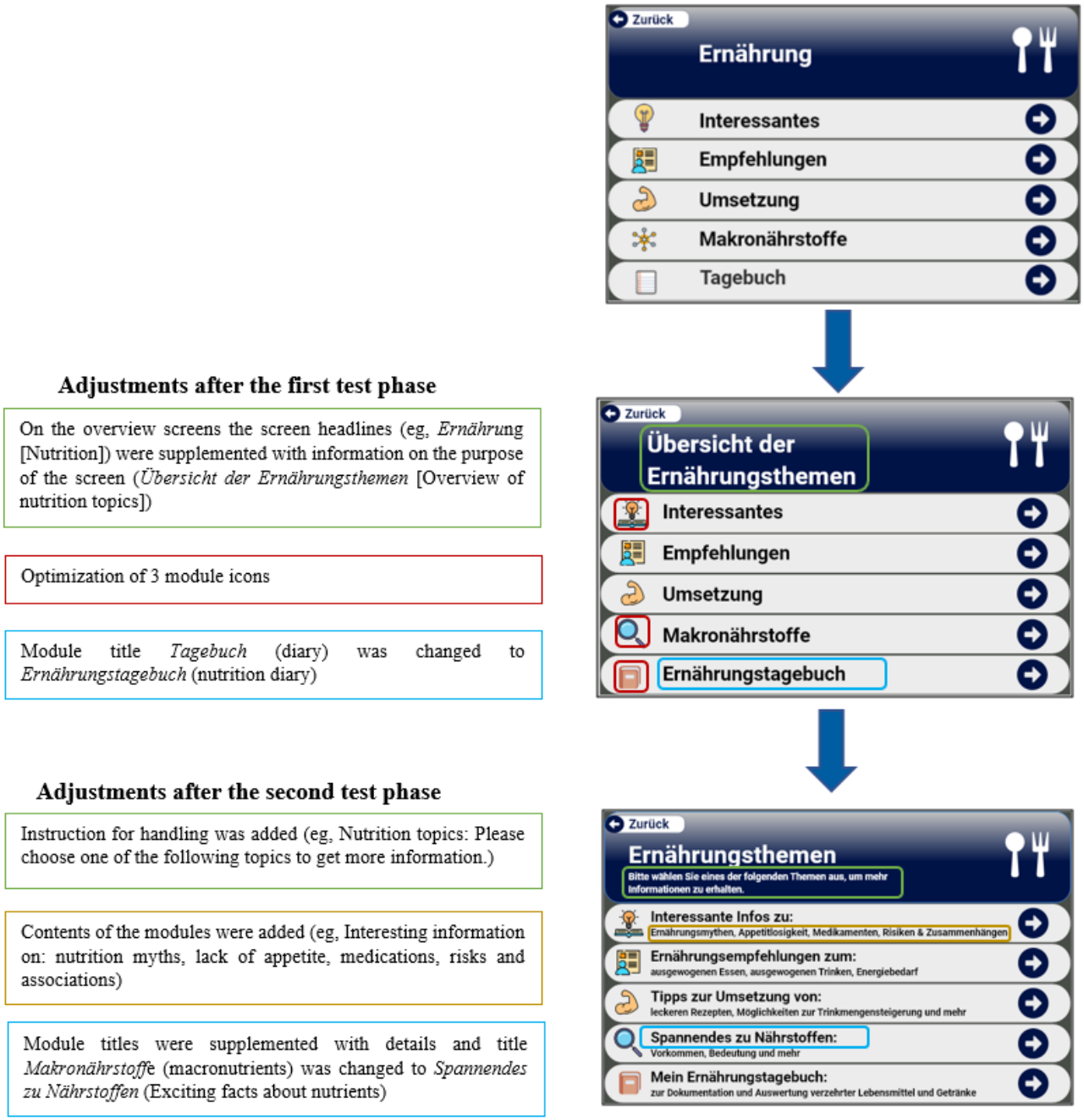
The tasks in the test sequence were always carried out in the same order, but patients had the option of skipping tasks at any time or stopping the usability test. The tests and verbal feedback during the tasks were recorded by taking notes.

After the usability test, patients were finally interviewed using the SUS. The questionnaire contained 10 statements about the usability of a system rated on a 5-point Likert scale. The values of individual items were added together and then multiplied by 2.5, resulting in a score between 0 and 100. The results provided a general overview of product usability [27].

Iterative Process

A total of 3 iterative test phases were used. Before each test phase, we refined the content, design, and potential functionality of the elements. The first test phase mainly tested the basic functionality of the chosen navigation structure with the target group, with *simpler* tasks such as finding the next page; in phases 2 and 3, we increased the number of navigation steps required in some tasks and tested other functions such as the use of the help button. To minimize the contact time and number of contacts with the study team in the context of the increasing incidence of COVID-19 in our region in Germany, it was decided to test the optimized elements and new content in the nutrition section only after completion of the second iterative test phase. As no further problems were found concerning the exercise part, but some were found concerning navigation from screens or the interpretation of nutritional diagrams, the last iterative phase focused on the questions from the nutrition part. Figure 2 shows an example of screenshots from version 1 to the final version. Multimedia Appendices 1 and 2 show click routes from the areas of nutrition and exercise, respectively.

Figure 2. German original version of screens from the 3 iterative test phases. Explanations and translations of changed elements based on the results of the test phase are shown in the text boxes.



Adjustments after the first test phase

On the overview screens the screen headlines (eg, *Ernährung* [Nutrition]) were supplemented with information on the purpose of the screen (*Übersicht der Ernährungsthemen* [Overview of nutrition topics])

Optimization of 3 module icons

Module title *Tagebuch* (diary) was changed to *Ernährungstagebuch* (nutrition diary)

Adjustments after the second test phase

Instruction for handling was added (eg, Nutrition topics: Please choose one of the following topics to get more information.)

Contents of the modules were added (eg, Interesting information on: nutrition myths, lack of appetite, medications, risks and associations)

Module titles were supplemented with details and title *Makronährstoffe* (macronutrients) was changed to *Spannendes zu Nährstoffen* (Exciting facts about nutrients)

Data Analysis

Statistical analyses were performed using the software SPSS (version 27.0; SPSS Inc).

Participants' characteristics, success rates of the tasks, and SUS scores were analyzed using descriptive statistics; they were presented as frequencies, means, SDs, and percentages.

The hints given during the usability test were presented as the total number of hints required for this task type in the iterative phase. The total number of unsolved tasks and percentages of unsolved tasks for all participants that performed this task type in the iterative phase were also reported.

The notes from the concurrent thinking aloud during the usability test were used to derive aspects that the participants noticed during the test. These aspects were discussed by the study team (LH and MS) to identify specific problems in the tasks and to derive possibilities for optimization.

An explorative ANOVA was conducted to compare the SUS scores among the 3 iterative test phases. The robust Welch *F* test was used if the assumption of homogeneity of variances was violated or if the data were not normally distributed.

The group was divided into 2 equally sized subgroups to test for group differences in successful task completion. Half of the participants who performed the tasks better were compared with the other half. The variables of age, sex, BMI, MNA-SF classification, technology affinity, TTM phase nutrition, TTM phase mobility, and SUS score were tested for group differences. The explorative analysis was performed using the

Mann–Whitney *U* test for ordinal and not normally distributed variables, and the chi-square test was performed for nominal variables. Statistical significance was set at $P<.05$, for all explorative analyses.

Results

Characteristics of Participants

A total of 49 patients who were aged 66-94 years (24/49, 49% female; mean 77.8, SD 6.2 years) participated in the study. Patient characteristics per iterative test phase are presented in Table 2. An example of screens and changed elements based on the results of the iterative test phases are shown in Figure 2.

Table 2. Overview of participants' characteristics within each iterative test phase (N=49).

Characteristics	Iterative phase 1		Iterative phase 2		Iterative phase 3[a]
	Nutrition	Mobility	Nutrition	Mobility	Nutrition
Participants[b]					
Total, n	15	12	13	13	12
Female, n (%)	8 (53)	6 (50)	9 (69)	5 (39)	5 (42)
BMI (kg/m^2), mean (SD)	26.3 (4.4)	26.6 (5.6)	27.6 (5.8)	26.6 (4.6)	27.6 (4.3)
Age (years), mean (SD)	79.1 (6.8)	78.4 (5.5)	78.5 (7.3)	76.3 (6.8)	76.8 (5.2)
TTM[c]phase, n (%)					
Precontemplation	3 (20)	1 (8)	4 (31)	0 (0)	3 (25)
Contemplation	2 (13)	5 (42)	3 (23)	5 (39)	2 (17)
Planning	4 (27)	3 (25)	1 (8)	4 (31)	5 (42)
Action	0 (0)	1 (8)	2 (15)	1 (8)	0 (0)
Maintenance	6 (40)	2 (17)	3 (23)	3 (23)	2 (17)
TC[d], mean (SD)					
Total score (12-60 points)	36.6 (11.3)	39.3 (13.4)	38.7 (11.4)	43.7 (6.9)	39.4 (11.9)
Points[e]	3.1 (1.0)	3.4 (1.6)	3.3 (1.1)	3.7 (0.6)	3.3 (1.0)

[a]Only tests in the nutrition section were performed to keep contact times and numbers low in the context of increasing COVID-19 incidence in the region.

[b]Different sample sizes within nutrition and mobility owing to the patient's choice option to participate only in one main theme or in both main themes.

[c]TTM: transtheoretical model of behavior change.

[d]TC: technology commitment (Neyer et al [42]).

[e]Average Likert scale points per item.

Navigation tasks required fewer hints (14.1%-26.5%) per task than comprehension tasks (20%-30.5%) in all iterative test phases (Table 3). The percentage of tasks that participants were unable to successfully complete was 0% for the navigation tasks in iterative test phase 1; 15.2% in iterative test phase 2; and 4.3% in iterative test phase 3. For the comprehension tasks, participants were unable to complete 1.9% of the total tasks in iterative test phase 1; 0% of the tasks in iterative test phase 2; and 2.5% of the tasks in iterative test phase 3 despite receiving hints. An overview of the tasks that could partially not be solved and the respective optimizations is provided in Multimedia Appendix 3.

Table 3. Participants' performance in the different navigation and comprehension tasks in the iterative test phases.

Tasks	Iterative phase 1			Iterative phase 2			Iterative phase 3		
	Total tasks[a], N	Hints[b], n (%)	Fail[c], n (%)	Total tasks, N	Hints, n (%)	Fail, n (%)	Total tasks, N	Hints, n (%)	Fail, n (%)
Navigation (total tasks)	177	25 (14)	0 (0)	204	54 (27)	31 (15)	93	16 (17)	4 (4)
Next screen (1 screen)	105	15 (14)	0 (0)	49	10 (20)	8 (16)	29	4 (14)	1 (4)
Further screen (≥2 screens)	21	3 (14)	0 (0)	54	20 (37)	10 (19)	37	8 (22)	3 (8)
Use of back button	32	6 (19)	0 (0)	55	10 (18)	6 (11)	18	1 (6)	0 (0)
Use of tab layout (text elements)	12	1 (8)	0 (0)	9	2 (22)	2 (22)	0	0 (0)	0 (0)
Use of the help button	0	0 (0)	0 (0)	28	10 (36)	5 (18)	9	3 (33)	1 (11)
Use of quizzes	7	0 (0)	0 (0)	0	0 (0)	0 (0)	0	0 (0)	0 (0)
Use of exercise diary	0	0 (0)	0 (0)	9	2 (22)	1 (11)	0	0 (0)	0 (0)
Comprehension (total tasks)	105	32 (31)	2 (2)	74	16 (22)	0 (0)	40	8 (20)	1 (3)
Purpose of screen	33	11 (33)	0 (0)	26	5 (19)	0 (0)	9	1 (11)	0 (0)
Foresight of content	15	6 (40)	0 (0)	0	0 (0)	0 (0)	12	2 (17)	0 (0)
Nutrition diary	8	6 (75)	2 (25)	0	0 (0)	0 (0)	0	0 (0)	0 (0)
Interpretation of content	38	7 (18)	0 (0)	72	11 (15)	0 (0)	19	5 (26)	1 (5)
Comprehension of quizzes	11	2 (19)	0 (0)	0	0 (0)	0 (0)	0	0 (0)	0 (0)

[a]Total number of tasks performed by all participants per iterative phase.

[b]Summed up the number and percentage of required hints for all participants for this task type in the iterative phase.

[c]Total number and percentage of unsolved tasks in all participants who performed this task type in the iterative phase.

Most participants understood how to use the buttons correctly. In 86.9% of all tests, participants consistently selected the arrow button for navigation to the next screen as intended. When navigating back to previous screens, this was done completely correctly in 84.7% of all tasks.

Many participants were able to successfully interpret the active input options such as the drinking protocol (11/15, 73%), the diagram of consumed food groups (9/10, 90%), and the exercise diary (7/8, 88%).

Some participants had difficulties in interpreting the content they would expect to find based on the names of the modules or the themes. The contents that needed to be optimized are shown in Table 4.

Many design elements on the overview screens and the educational features were rated positively by older adults. For educational content with texts, the font sizes, readability, and length of the text were positively rated in almost all corresponding tasks (92.3%).

Moreover, more than half of the participants (25/49, 51%) were able to successfully solve more than 90% of the tasks. Exploratory analysis of group differences suggested that those who solved more than 90% of the tasks had significantly higher technology affinity (P=.02). The participants who were able to solve more than 90% of the tasks also rated the e-coach significantly better with the SUS score (P=.04).

Table 4. Content and structures optimized following iterative testing with participants.

Domain	Adaptations
Navigation	• On the main overview screens for the nutrition and mobility modules, details on the content of modules were added. • Checkboxes for the confirmation of exercise execution and the labeling of the elements were enlarged. • For screens that guide different topics in a module, a question or more guidance about the content was added in addition to the title (eg, increasing activity: How can I become more active in everyday life?).
Comprehension	• Keywords were supplemented with further information (eg, *nutrition* was changed to *nutrition topics*). • The wording *macronutrients* was changed to *nutrients*. • An instruction for the action was added below the screen heading (eg, "Please select one of the following topics to get more information."). • Information for food groups was added (*2/5* was changed to *2 of 5 servings*).
Design	• The symbols for nutrients (the molecule symbol was changed to a magnifying glass), interesting information (the light bulb was changed to a book with light bulb on it), and the nutrition diary (the booklet was changed to a book) were replaced. • Photographs for text elements were exchanged for symbols or drawings. • Exercise photos were used instead of exercise drawings; a white background was added to the exercise photos. • Any other elements besides diagrams were removed from the evaluation screens. • Feedback on reaching the training goal using flowers instead of stars was added (flowers contain additional information about the number of exercises performed).

Evaluation of the SUS

The evaluation of the SUS showed a continuous improvement in the usability of the e-coach (Table 5).

Because the normality assumptions of the ANOVA were violated, a 1-way Welch-ANOVA was performed to determine whether the SUS score was significantly different among the test phases. The improvement in the SUS score was not statistically significant among the 3 test phases (Welch $F_{2,46}$=1.79; P=.19).

Table 5. System Usability Scale (SUS) score for each iterative test phase.

Phase	Values		
	n (%)	Mean (SD; range)	Median (IQR)
Iterative phase 1	21 (43)	69.3 (16.3; 42.5-97.5)	65.0 (57.5-83.8)
Iterative phase 2	16 (33)	70.3 (18.7; 20.0-95.0)	77.5 (60.0-84.4)
Iterative phase 3	12 (25)	78.1 (11.8; 50.0-92.5)	82.5 (70.6-86.9)

Discussion

Principal Findings

We showed that it is possible to conduct iterative test phases and improve the usability of a health app, even for older adults with health restrictions who are undergoing inpatient rehabilitation. After 3 iterative test phases, the average SUS increased from 69.3 (SD 16.3; median 65) to 78.1 (SD 11.8; median 82.5), indicating good usability of the e-coach and is comparable with the SUS results on eHealth application use in other studies with older adults [45-47]. More hints were required to answer comprehension questions (20%-30.5%) than to solve navigation tasks (14.1%-26.5%) during testing. With minor support from the study team, by hints, it was almost always possible for participants to solve the tasks or questions. Overall, only 5.5% of all tasks in all tests could not be successfully completed. These results show that even participants who had greater difficulties using the e-coach were able to solve most tasks with minor support.

Comparing the completion of tasks among the 3 iterations, it is noticeable that in the first iterative test phase, hints were needed, especially for comprehension questions (30.5%), and a few tasks were not fulfilled by all participants in this test (1.9%). In this iterative phase, hints were often needed for comprehension questions, such as what can be done on certain screens and also concerning the interpretation of what was meant by certain titles (eg, an overview of the recommended amount of physical activity). These difficulties were solved by adding more information to the descriptive texts and partly by changing the wording of individual terms.

In the second iterative test phase, the proportion of required hints in the navigation tasks (26.5%) increased noticeably compared with the first iterative test phase (14.1%), and 15.2% of all navigation tasks could not be solved by the participants in the second iterative test phase. In the first iterative test phase, many navigation tasks included easier functions, such as navigation to the next screen, and only a few tasks with more complex navigation steps. In contrast, in the second iterative test phase, more complex tasks were added, requiring, for

example, 2 navigation steps, and the proportion of simpler navigation tasks decreased accordingly (Table 1). It seems possible that the increase in the complexity of the navigation tasks explains the increase in the number of hints required, as well as the increase in the proportion of navigation tasks that could not be solved. However, we did not perform any corresponding measurements that would allow conclusions to be drawn about the participants' working memory capacity, as this was not part of our study.

In the third iterative test phase, hints were also required more often for navigation tasks that required several navigation steps. Although there were some tasks in the last iteration phase that required more than one navigation step, only 17% (2/12) of all participants needed help with all navigation tasks, and 4.3% of all navigation tasks in the third iterative test phase were not successfully completed. However, in the comprehension tasks, the need for hints was almost similar in the last iterative phase (20%) compared with the second iterative phase (21.6%), but more tasks could not be solved (2.5%) than in the second iterative phase. Compared with the first iterative test phase, almost all participants (10/12, 83%) in the last iterative test phase were able to derive the following content from the labeling of the modules. In addition, the question about what can be done on the overview screens was also answered correctly by all participants. Nevertheless, some participants found it difficult to interpret diagrams, meaning that they required hints, and one participant was unable to solve the task.

Other aspects should be considered as possible reasons for the different performance in the tasks, in addition to the increase in complexity. There were fewer tasks and fewer participants in the third iterative test phase because increasing COVID-19 case numbers forced us to reduce contact times with the participants. In addition, all the comprehension questions that could not be answered successfully in the last test were questions about protein and energy intake shown in a diagram. This topic could be difficult for people with little existing knowledge of the subject.

Furthermore, in our exploratory analysis of group differences between half of the participants who solved more than 90% of the tasks correctly and those who solved less than 90% of the tasks correctly, we found indications that those who solved more than 90% of the tasks correctly had a significantly higher affinity for technology (P=.02). A significant correlation between the use of health apps by older people and a higher affinity for technology has also been found in other studies [9,48]. Moreover, a stronger interest in technology appears to have a significant influence on the use of information and communication technologies even among people aged >80 years [49]. To estimate how well older adults can cope with a health app or whether more support might be needed, technology affinity could be surveyed beforehand and used as an indicator of the need for support. Different variables such as education level, sex, and computer literacy did not significantly influence the use of health apps in the study by Rasche et al [9]. Age differed significantly among groups. Individuals who reported using health apps were significantly younger than those who did not use health apps. In contrast, Schlomann et al [49] also found no significant influence of the variable age among older adults who already used mobile devices (smartphones, tablets, fitness trackers, or smartwatches). In our study, individuals who correctly solved more than 90% of the tasks were younger (mean 76.0, SD 4.3 years) than those who correctly solved less than 90% of the tasks (mean 79.7, SD 7.4 years) but this difference was not statistically significant (P=.09).

Technology commitment in our study population was at a median of 3.5 (SD 0.6) points in the group of participants who correctly solved more than 90% of the tasks and at a median of 3.0 (0.6) in the group of people who correctly solved less than 90% of the tasks. This value seems to correspond approximately to the technology readiness of people in this age group in Germany. Rasche et al [9] found comparable values among older people who reported no app use (mean 2.9, SD 0.6 points) or app use but no health app use (mean 3.4, SD 0.6 points).

To increase the usability of e-coaches in health care, as well as for less technically inclined older persons, detailed instructions on how to use the e-coach could be applied. Older people seem to have greater benefits from step-by-step instructions when learning to use new technologies. To address this problem in a real-world setting in terms of use of the e-coach by older adults, guidance (eg, in the form of a printed manual) should be additionally offered [50]. Furthermore, it was helpful for participants if the screens contained additional information, for example, what should be done on the screen or what content and information they would receive if they selected certain modules. The screens were not rated as overly cluttered despite additional information. This may be an important fact for other researchers developing apps for multi-morbid older adults in rehabilitation facilities, who have rarely been involved in app development to this extent. Although studies on eHealth interventions are already being conducted in the context of geriatric rehabilitation, apps have only been considered as a form of intervention in a few studies to date, and many applications relate to specific clinical conditions, particularly in the neurological field [14]. In older adults with limited mobility and after hip fractures, positive effects on mobility, functional outcomes, and hospitalization could be achieved through telerehabilitation interventions with an app [16] or via web sites [15,51]. However, these applications do not focus on educating participants but rather on teaching exercises. People who have not yet reached this phase in their behavioral change process may therefore benefit less from these tools. Furthermore, it is possible to use apps to conduct relevant assessments for older people with multiple morbidities with regard to fall risk [46] and mobility [47]. The usability of these assessment apps was investigated in older people and achieved SUS scores of 77.9 and 77.6, comparable with the score gained by our e-coach. In general, there are already different apps in the rehabilitation context for postoperative care after certain surgical interventions [52] or for improving self-management of hypertension or for medication planning. In these studies, however, it became clear that apps from the app store often lacked an evidence-based background, usability for older people is poor, and there is also a lack of data security [53,54]. In the field of nutrition apps, no studies involving participants aged >70 years have been found in a review from 2019 [18].

The data on the TTM phases of the participants demonstrate that geriatric patients in rehabilitation are at different phases of the behavioral change process. In the iterative tests with nutrition content as well as with physical activity content, the proportion of participants in the first 2 phases of TTM (precontemplation and contemplation) and the remaining phases (preparation, action, and maintenance) was quite balanced. As such, the readiness and implementation of behavior change in the areas of nutrition and physical activity appear to be heterogeneous among the target group, requiring the use of different strategies in the e-coach to support behavior changes. No indications of significant group differences between the half of the participants who were able to solve more than 90% of their tasks and those who were able to solve less than 90% of the tasks were found for the TTM phase distribution in the area of nutrition. For the TTM phase in the physical activity domain, no testing for group differences was conducted because of the small sample size resulting from missing data from the last iterative test. A correlation between certain phases of behavior change and the ability to successfully solve tasks could be explained only to a limited extent. In the case of tasks for operating the app, such as finding specific screens or using buttons, personal readiness to change one's nutritional or physical activity behavior should have little influence. It could be possible that the interpretation of specific content with terms such as *protein* or *foodgroup* is more difficult for people with less existing knowledge. However, even someone without existing technical knowledge may have already achieved the nutrition and physical activity goals we used. Therefore, it is not practical to draw conclusions from the behavior change phase alone about the ability to use or the general understanding of the app based on our results.

When developing the e-coach, design recommendations for the target group [29,30] were considered in advance. The button labels and the texts on the navigation screens were phrased without technical terms to make them usable even for novice technology users. This, as well as the chosen high contrasts of the different contents and texts, was maintained throughout the development process. In some cases, the minimum size of the arrow buttons was exceeded, but based on the feedback from our tests, this did not represent a barrier for the patients.

In a few tests, it was noticed that a user wanted to select the correct button, but the pressure on the button was not recognized by the device. Besides a possible malfunction of the touchscreen, there are 2 points that have been described in further studies [30,55] and are also found in our study; these should also be considered with regard to other apps for this target group. One possible cause of failure could be the contact between the test person's finger and the touchscreen. In addition, the conductivity of the skin decreases with age, but we also observed that some participants typed using their fingernails rather than their fingertips. To improve usability, older participants could alternatively be offered to operate the device with a stylus [30]. As a second possible source of error, we observed the button being pressed for a very long time or the finger already being positioned outside the button on the touchscreen and then moved toward the button. When training users, the use of buttons should be explained in more detail and practiced. Furthermore, feedback (haptic, auditory, or visual) could improve usability [30]; however, this additional function was not used in this study.

Limitations

It is likely that more people who already had a general interest in nutrition, physical activity, or technology participated in the study. Therefore, a selection bias cannot be ruled out even if at least the measured technology commitment also corresponds to the figures from another study with older people in Germany with varying degrees of experience in the use of technology [9]. In addition, study participation was not dependent on specific criteria related to mobility limitation or malnutrition. In a recent review with meta-analysis on nutritional status and physical functionality of geriatric patients in rehabilitation, it was demonstrated that a large proportion of persons in geriatric rehabilitation are affected by malnutrition and mobility restrictions [3]. The results of our tests also confirmed this finding for the participants of this study. Malnutrition was present in 27% (13/49) of the participants, and the risk of malnutrition was present in 51% (25/49) of the participants, as measured by the MNA-SF. The PASE score was 46.2, which is significantly lower than the average score of community-dwelling older adults reported by Washburn et al [56].

Data collection was conducted by study team members who were involved in the development of the e-coach. An uninvolved person would have had to conduct usability tests to increase the objectivity of these tests. This was not possible because of the financial resources of the project.

Moreover, the test situation itself may have biased the results. On the one hand, it is likely that older adults were more nervous and made more mistakes than if they had used the e-coach unobserved. On the other hand, the possibility of receiving hints and help from the examiner could also have led to a situation in which help was requested more quickly and, if necessary, tasks would have been solved after some time without the help of the study team member. In addition, it cannot be ruled out that unintentional nonverbal responses were given by the study team (eg, nodding) that were not recorded in the documentation.

In the context of this study, the time spent on the task was not determined. The time it takes a participant to complete a task can be an indicator of usability, as quick performance can also indicate ease of use. However, there is also evidence from a study by Sonderegger et al [57] that the time to complete a task is less related to perceived usability in older people than in younger people. We decided not to measure the time during tasks so as not to create additional pressure in the already unfamiliar test situation and because time-critical aspects presumably tend to play a minor role in the use by our target group.

This study only evaluated whether e-coach elements are generally usable for older adults. It provides an indication of the e-coach's usability for first-time users but not for a longer period of use.

Conclusions

This study involved older people undergoing inpatient rehabilitation in iterative optimization and usability testing for an e-coach according to the German International Organization for Standardization 9241-210:2019 *Ergonomics of human–system interaction—Part 210: Human-centered design of interactive systems* [24] and the user-centered design process [23]. It has been shown that this approach can be successfully applied to this vulnerable and low technologically skilled target group. The involvement of the target group was very important in developing a program that older people could rely on that is oriented to their needs, that is based on a psychological model for long-term behavior change, and that can also be used by them. When an app addresses important health issues (eg, malnutrition and inactivity), it seems particularly important to consider known barriers such as a lack of confidence in the app, low usability, or even users' low existing technical experience and to offer evidence-based support [9].

As the target group is particularly vulnerable, and an individual's willingness to continuously use the e-coach may impose an additional burden on them, it is essential to evaluate the acceptance, willingness, and adherence for long-term use of the system in a further study. Previous studies have shown that it can be very difficult to recruit patients for such long-term use of technical devices in this target group, and that good usability, as well as subjective benefits for patients, must necessarily be present [58].

Acknowledgments

We thank all participants for participating in the study and for providing us with valuable inputs. Special thanks are due to Daniel Küppers for programing the e-coach and Patrick Elfert for his valuable input and work in developing the nutrition diary. We would also like to thank the geriatric and cardiologic rehabilitation departments of the Rehabilitation Center of Oldenburg for giving us the opportunity to conduct our study.

This research was funded by an intramural grant from Carl von Ossietzky University Oldenburg (School of Medicine and Health Sciences; Research Pool 2018-018). The funder had no role in the study design, data collection, analysis and interpretation, decision to publish, or preparation of the manuscript.

Authors' Contributions

LH, AH, and RD led to the conception and design of the study. LH, MS, AH, and RD contributed to their expertise in developing the app content. LH and MS coordinated and performed usability testing and evaluation. LH wrote the manuscript with input from MS, AH, and RD. AH and RD obtained funding and supervised the study. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of tasks in iterative phase 2 physical activity.
[PNG File , 1058 KB - humanfactors_v9i1e31823_app1.png]

Multimedia Appendix 2
Examples of tasks in iterative phase 3 nutrition.
[PNG File , 1151 KB - humanfactors_v9i1e31823_app2.png]

Multimedia Appendix 3
Unsolved tasks and optimizations.
[PDF File (Adobe PDF File), 722 KB - humanfactors_v9i1e31823_app3.pdf]

References

1. Benzinger P, Eidam A, Bauer JM. [Clinical importance of the detection of frailty]. Z Gerontol Geriatr 2021 May;54(3):285-296 [FREE Full text] [doi: 10.1007/s00391-021-01873-z] [Medline: 33782735]
2. Nascimento CM, Ingles M, Salvador-Pascual A, Cominetti MR, Gomez-Cabrera MC, Viña J. Sarcopenia, frailty and their prevention by exercise. Free Radic Biol Med 2019 Feb 20;132:42-49. [doi: 10.1016/j.freeradbiomed.2018.08.035] [Medline: 30176345]
3. Wojzischke J, van Wijngaarden J, van den Berg C, Cetinyurek-Yavuz A, Diekmann R, Luiking Y, et al. Nutritional status and functionality in geriatric rehabilitation patients: a systematic review and meta-analysis. Eur Geriatr Med 2020 Apr;11(2):195-207. [doi: 10.1007/s41999-020-00294-2] [Medline: 32297199]

4. van Balen R, Gordon AL, Schols JM, Drewes YM, Achterberg WP. What is geriatric rehabilitation and how should it be organized? A Delphi study aimed at reaching European consensus. Eur Geriatr Med 2019 Dec;10(6):977-987. [doi: 10.1007/s41999-019-00244-7] [Medline: 34652774]
5. Grund S, van Wijngaarden JP, Gordon AL, Schols JM, Bauer JM. EuGMS survey on structures of geriatric rehabilitation across Europe. Eur Geriatr Med 2020 Apr;11(2):217-232. [doi: 10.1007/s41999-019-00273-2] [Medline: 32297190]
6. Hill A, Hoffmann T, McPhail S, Beer C, Hill KD, Brauer SG, et al. Factors associated with older patients' engagement in exercise after hospital discharge. Arch Phys Med Rehabil 2011 Sep;92(9):1395-1403. [doi: 10.1016/j.apmr.2011.04.009] [Medline: 21878210]
7. Forkan R, Pumper B, Smyth N, Wirkkala H, Ciol MA, Shumway-Cook A. Exercise adherence following physical therapy intervention in older adults with impaired balance. Phys Ther 2006 Mar;86(3):401-410. [Medline: 16506876]
8. Bitkom. Welche der folgenden Gesundheits-Apps nutzen Sie bereits bzw. können Sie sich vorstellen, zukünftig zu nutzen? Statista. 2020 Sep. URL: https://de.statista.com/statistik/daten/studie/1173923/umfrage/nutzung-von-gesundheits-apps-durch-senioren/ [accessed 2021-04-06]
9. Rasche P, Wille M, Bröhl C, Theis S, Schäfer K, Knobe M, et al. Prevalence of health app use among older adults in Germany: national survey. JMIR Mhealth Uhealth 2018 Jan 23;6(1):e26 [FREE Full text] [doi: 10.2196/mhealth.8619] [Medline: 29362211]
10. Seifert A, Meidert U. Quantified seniors. Technically assisted self-measurement among older adults. Präv Gesundheitsf 2018 Apr 12;13(4):353-360. [doi: 10.1007/s11553-018-0646-1]
11. Mehra S, Visser B, Dadema T, van den Helder J, Engelbert RH, Weijs PJ, et al. Translating behavior change principles into a blended exercise intervention for older adults: design study. JMIR Res Protoc 2018 May 02;7(5):e117 [FREE Full text] [doi: 10.2196/resprot.9244] [Medline: 29720358]
12. Bickmore TW, Silliman RA, Nelson K, Cheng DM, Winter M, Henault L, et al. A randomized controlled trial of an automated exercise coach for older adults. J Am Geriatr Soc 2013 Oct;61(10):1676-1683. [doi: 10.1111/jgs.12449] [Medline: 24001030]
13. Silveira P, van de Langenberg R, van Het Reve E, Daniel F, Casati F, de Bruin ED. Tablet-based strength-balance training to motivate and improve adherence to exercise in independently living older people: a phase II preclinical exploratory trial. J Med Internet Res 2013 Aug 12;15(8):e159 [FREE Full text] [doi: 10.2196/jmir.2579] [Medline: 23939401]
14. Kraaijkamp JJ, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. J Med Internet Res 2021 Aug 19;23(8):e24015 [FREE Full text] [doi: 10.2196/24015] [Medline: 34420918]
15. Ortiz-Piña M, Molina-Garcia P, Femia P, Ashe MC, Martín-Martín L, Salazar-Graván S, et al. Effects of tele-rehabilitation compared with home-based in-person rehabilitation for older adult's function after hip fracture. Int J Environ Res Public Health 2021 May 20;18(10):5493 [FREE Full text] [doi: 10.3390/ijerph18105493] [Medline: 34065523]
16. Bean JF, Brown L, DeAngelis TR, Ellis T, Kumar VS, Latham NK, et al. The rehabilitation enhancing aging through connected health prehabilitation trial. Arch Phys Med Rehabil 2019 Nov;100(11):1999-2005. [doi: 10.1016/j.apmr.2019.04.015] [Medline: 31152705]
17. Best R, Souders DJ, Charness N, Mitzner TL, Rogers WA. The role of health status in older adults' perceptions of the usefulness of eHealth technology. Hum Asp IT Aged Popul (2015) 2015 Sep;9194:3-14 [FREE Full text] [doi: 10.1007/978-3-319-20913-5_1] [Medline: 31240277]
18. Villinger K, Wahl DR, Boeing H, Schupp HT, Renner B. The effectiveness of app-based mobile interventions on nutrition behaviours and nutrition-related health outcomes: a systematic review and meta-analysis. Obes Rev 2019 Oct;20(10):1465-1484 [FREE Full text] [doi: 10.1111/obr.12903] [Medline: 31353783]
19. Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (ISO 9241-210:2019); German version EN ISO 9241-210:2019. Beuth Publishing DIN. 2020 Mar. URL: https://www.beuth.de/de/norm/din-en-iso-9241-210/313017070 [accessed 2022-03-10]
20. Richter M, Flückiger M. Usability Engineering Kompakt: Benutzbare Produkte Gezielt Entwickeln. Berlin, Heidelberg: Springer; 2013:179.
21. Happe L, Hein A, Diekmann R. What do geriatric rehabilitation patients and experts consider relevant? Requirements for a digitalised e-coach for sustainable improvement of nutrition and physical activity in older adults - a qualitative focus group study. BMC Geriatr 2021 Dec 18;21(1):712 [FREE Full text] [doi: 10.1186/s12877-021-02692-y] [Medline: 34922492]
22. Warschburger P. In: Warschburger P, editor. Beratungspsychologie. Berlin, Heidelberg: Springer; 2009:81-103.
23. Romain A, Bortolon C, Gourlan M, Carayol M, Decker E, Lareyre O, et al. Matched or nonmatched interventions based on the transtheoretical model to promote physical activity. A meta-analysis of randomized controlled trials. J Sport Health Sci 2018 Jan;7(1):50-57 [FREE Full text] [doi: 10.1016/j.jshs.2016.10.007] [Medline: 30356472]
24. DiClemente CC, Prochaska JO, Gibertini M. Self-efficacy and the stages of self-change of smoking. Cogn Ther Res 1985 Apr;9(2):181-200. [doi: 10.1007/bf01204849]
25. Prochaska JO, Butterworth S, Redding CA, Burden V, Perrin N, Leo M, et al. Initial efficacy of MI, TTM tailoring and HRI's with multiple behaviors for employee health promotion. Prev Med 2008 Mar;46(3):226-231 [FREE Full text] [doi: 10.1016/j.ypmed.2007.11.007] [Medline: 18155287]

26. Jiménez-Zazo F, Romero-Blanco C, Castro-Lemus N, Dorado-Suárez A, Aznar S. Transtheoretical model for physical activity in older adults: systematic review. Int J Environ Res Public Health 2020 Dec 11;17(24):9262 [FREE Full text] [doi: 10.3390/ijerph17249262] [Medline: 33322327]
27. Brooke J. SUS: A 'Quick and Dirty' usability scale. In: Jordan PW, Weerdmeester B, Thomas A, McIelland IL, editors. Usability Evaluation in Industry. London: Taylor and Francis; 1996:189-194.
28. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behavior Research Methods, Instruments, & Computers 2003 Aug;35(3):379-383. [doi: 10.3758/BF03195514]
29. Eichhorn C, Plecher D, Lurz M, Leipold N, Böhm M, Krcmar H, et al. Combining motivating strategies with design concepts for mobile apps to increase usability for the elderly and Alzheimer patients. In: HCII 2020: Human Aspects of IT for the Aged Population. Healthy and Active Aging. Cham: Springer International Publishing; 2020 Jul Presented at: International Conference on Human-Computer Interaction; Jul 19-24, 2020; Copenhagen, Denmark p. 47-66. [doi: 10.1007/978-3-030-50249-2_4]
30. Nurgalieva L, Laconich JJ, Baez M, Casati F, Marchese M. A systematic literature review of research-derived touchscreen design guidelines for older adults. IEEE Access 2019;7:22035-22058. [doi: 10.1109/access.2019.2898467]
31. Volkert D, Beck AM, Cederholm T, Cruz-Jentoft A, Goisser S, Hooper L, et al. ESPEN guideline on clinical nutrition and hydration in geriatrics. Clin Nutr 2019 Feb;38(1):10-47. [doi: 10.1016/j.clnu.2018.05.024] [Medline: 30005900]
32. Deutsche Gesellschaft für Ernährung e. V.. DGE Praxiswissen: Essen und Trinken im Alter. In: Fit Im Alter: Gesund Essen, Besser Leben. Bonn: Deutsche Gesellschaft für Ernährung; 2014:34.
33. Deutsche Gesellschaft für Ernährung e.V.. DGE Praxiswissen. Mangelernährung im Alter. In: Fit Im Alter: Gesund Essen, Besser Leben. Bonn: Deutsche Gesellschaft für Ernährung; 2012:35.
34. Bundeszentrale für gesundheitliche Aufklärung. Aktiv im Alltag, aktiv im Leben: Fit und eigenständig bleiben: Anregungen für Menschen mit Bewegungseinschränkungen. In: Älter Werden in Balance. Köln: Bundeszentrale für gesundheitliche Aufklärung; May 2015.
35. Liu-Ambrose T, Donaldson MG, Ahamed Y, Graf P, Cook WL, Close J, et al. Otago home-based strength and balance retraining improves executive functioning in older fallers: a randomized controlled trial. J Am Geriatr Soc 2008 Oct;56(10):1821-1830. [doi: 10.1111/j.1532-5415.2008.01931.x] [Medline: 18795987]
36. Clemson L, Singh MA, Bundy A, Cumming RG, Manollaras K, O'Loughlin P, et al. Integration of balance and strength training into daily life activity to reduce rate of falls in older people (the LiFE study): randomised parallel trial. Br Med J 2012 Aug 07;345:e4547 [FREE Full text] [doi: 10.1136/bmj.e4547] [Medline: 22872695]
37. Clegg A, Barber S, Young J, Iliffe S, Forster A. The Home-based Older People's Exercise (HOPE) trial: a pilot randomised controlled trial of a home-based exercise intervention for older people with frailty. Age Ageing 2014 Sep;43(5):687-695 [FREE Full text] [doi: 10.1093/ageing/afu033] [Medline: 24742587]
38. Kaiser MJ, Bauer JM, Ramsch C, Uter W, Guigoz Y, Cederholm T, MNA-International Group. Validation of the Mini Nutritional Assessment short-form (MNA-SF): a practical tool for identification of nutritional status. J Nutr Health Aging 2009 Nov;13(9):782-788. [doi: 10.1007/s12603-009-0214-7] [Medline: 19812868]
39. Deutsche Gesellschaft für Ernährung e. V.. Dge-qualitätsstandard Für Die Verpflegung Mit Essen Auf Rädern Und in Senioreneinrichtungen. Bonn: Deutsche Gesellschaft für Ernährung; 2020.
40. Washburn R, Smith K, Jette A, Janney C. The physical activity scale for the elderly (PASE): development and evaluation. J Clin Epidemiol 1993 Feb;46(2):153-162 [FREE Full text] [doi: 10.1016/0895-4356(93)90053-4]
41. Pfeifer K, Rütten A. [National recommendations for physical activity and physical activity promotion]. Gesundheitswesen 2017 Mar;79(S 01):2-3. [doi: 10.1055/s-0042-123346] [Medline: 28399579]
42. Neyer FJ, Felber J, Gebhardt C. Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft. Diagnostica 2012 Apr;58(2):87-99. [doi: 10.1026/0012-1924/a000067]
43. Silva P, Nunes F. 3 x 7 usability testing guidelines for older adults. In: Proceedings of the 3rd Mexican Workshop on Human Computer Interaction (MexIHC'2010). 2010 Presented at: 3rd Mexican Workshop on Human Computer Interaction (MexIHC'2010); Nov 8-10, 2010; San Luis Potosí, México URL: https://mural.maynoothuniversity.ie/6030/1/PS-Usability-Testing.pdf
44. Ericsson K, Simon HA. Protocol Analysis : Verbal Reports as Data. Cambridge, MA: MIT Press; 1984.
45. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usab Stud Arch 2009:114-123 [FREE Full text]
46. Hsieh KL, Fanning JT, Rogers WA, Wood TA, Sosnoff JJ. A fall risk mHealth app for older adults: development and usability study. JMIR Aging 2018 Nov 20;1(2):e11569 [FREE Full text] [doi: 10.2196/11569] [Medline: 31518234]
47. Bergquist R, Vereijken B, Mellone S, Corzani M, Helbostad JL, Taraldsen K. App-based self-administrable clinical tests of physical function: development and usability study. JMIR Mhealth Uhealth 2020 Apr 27;8(4):e16507 [FREE Full text] [doi: 10.2196/16507] [Medline: 32338616]
48. Seifert A, Vandelanotte C. The use of wearables and health apps and the willingness to share self-collected data among older adults. Aging Health Res 2021 Sep;1(3):100032 [FREE Full text] [doi: 10.1016/j.ahr.2021.100032]

49. Schlomann A, Seifert A, Zank S, Rietz C. Assistive technology and mobile ICT usage among oldest-old cohorts: comparison of the oldest-old in private homes and in long-term care facilities. Res Aging 2020;42(5-6):163-173. [doi: 10.1177/0164027520911286] [Medline: 32167019]
50. Hickman J, Rogers W, Fisk A. Training older adults to use new technology. J Gerontol B Psychol Sci Soc Sci 2007 Jun;62 Spec No 1:77-84. [doi: 10.1093/geronb/62.special_issue_1.77] [Medline: 17565168]
51. Báez M, Ibarra F, Far I, Ferron M, Casati F. Online group-exercises for older adults of different physical abilities. In: Proceedings of the International Conference on Collaboration Technologies and Systems (CTS). 2016 Presented at: International Conference on Collaboration Technologies and Systems (CTS); Oct. 31 - Nov. 4, 2016; Orlando, FL, USA. [doi: 10.1109/cts.2016.0098]
52. Patel B, Thind A. Usability of mobile health apps for postoperative care: systematic review. JMIR Perioper Med 2020 Jul 20;3(2):e19099 [FREE Full text] [doi: 10.2196/19099] [Medline: 33393925]
53. Alessa T, Hawley MS, Hock ES, de Witte L. Smartphone apps to support self-management of hypertension: review and content analysis. JMIR Mhealth Uhealth 2019 May 28;7(5):e13645 [FREE Full text] [doi: 10.2196/13645] [Medline: 31140434]
54. Stuck RE, Chong AW, Mitzner TL, Rogers WA. Medication management apps: usable by older adults? Proc Hum Factors Ergon Soc Annu Meet 2017 Sep;61(1):1141-1144 [FREE Full text] [doi: 10.1177/1541931213601769] [Medline: 29158662]
55. Kobayashi M, Hiyama A, Miura T, Asakawa C, Hirose M, Ifukube T. Elderly user evaluation of mobile touchscreen interactions. In: Campos P, Graham N, Jorge J, Nunes N, Palanque P, Winkler M, editors. Human-Computer Interaction – INTERACT 2011. Berlin, Heidelberg: Springer; 2011:83-99.
56. Washburn RA, McAuley E, Katula J, Mihalko SL, Boileau RA. The physical activity scale for the elderly (PASE): evidence for validity. J Clin Epidemiol 1999 Jul;52(7):643-651. [doi: 10.1016/s0895-4356(99)00049-9] [Medline: 10391658]
57. Sonderegger A, Schmutz S, Sauer J. The influence of age in usability testing. Appl Ergon 2016 Jan;52:291-300. [doi: 10.1016/j.apergo.2015.06.012] [Medline: 26360221]
58. Nymberg VM, Bolmsjö BB, Wolff M, Calling S, Gerward S, Sandberg M. 'Having to learn this so late in our lives…' Swedish elderly patients' beliefs, experiences, attitudes and expectations of e-health in primary health care. Scand J Prim Health Care 2019 Mar;37(1):41-52 [FREE Full text] [doi: 10.1080/02813432.2019.1570612] [Medline: 30732519]

Abbreviations

MNA-SF: Mini Nutritional Assessment-Short Form
PASE: Physical Activity Scale for the Elderly
SUS: System Usability Scale
TTM: transtheoretical model of behavior change

Edited by A Kushniruk; submitted 12.07.21; peer-reviewed by S Mehra, P Rasche, E Stringer; comments to author 06.09.21; revised version received 01.11.21; accepted 29.12.21; published 16.03.22.

Please cite as:
Happe L, Sgraja M, Hein A, Diekmann R
Iterative Development and Applicability of a Tablet-Based e-Coach for Older Adults in Rehabilitation Units to Improve Nutrition and Physical Activity: Usability Study
JMIR Hum Factors 2022;9(1):e31823
URL: https://humanfactors.jmir.org/2022/1/e31823
doi:10.2196/31823
PMID:35293874

Original Paper

An Interactive Voice Response System to Increase Physical Activity and Prevent Cancer in the Rural Alabama Black Belt: Design and Usability Study

Mohanraj Thirumalai[1], MS, MEng, PhD; Nashira Brown[2], MS; Soumya Niranjan[1], BPharm, MS, PhD; Sh'Nese Townsend[3], MHA; Mary Anne Powell[3]; Whitney Neal[3], MA; Erica Schleicher[3]; Venkatadri Raparla[4], MS; Robert Oster[5], PhD; Wendy Demark-Wahnefried[2], PhD; Dori Pekmezi[3], PhD

[1]Department of Health Services Administration, School of Health Professions, University of Birmingham at Alabama, Birmingham, AL, United States

[2]Department of Nutrition Sciences, School of Health Professions, University of Birmingham at Alabama, Birmingham, AL, United States

[3]Department of Health Behavior, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, United States

[4]University of Alabama at Birmingham/Lakeshore Research Collaborative, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States

[5]Division of Preventive Medicine, Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, AL, United States

Corresponding Author:
Mohanraj Thirumalai, MS, MEng, PhD
Department of Health Services Administration
School of Health Professions
University of Birmingham at Alabama
SHPB #590E
1716 9th Avenue South
Birmingham, AL, 35244
United States
Phone: 1 2059347189
Email: mohanraj@uab.edu

Abstract

Background: Increased physical activity (PA) levels are associated with reduced risk and improved survival for several cancers; however, most Americans engage in less than the recommended levels of PA. Using interactive voice response (IVR) systems to provide personalized health education and counseling may represent a high-reach, low-cost strategy for addressing physical inactivity and cancer disparities in disproportionately burdened rural regions. However, there has been a paucity of research conducted in this area to date.

Objective: The aim of this study is to design, develop, and test the usability of an IVR system aimed at increasing PA levels in the rural Alabama Black Belt.

Methods: A pilot version of the IVR system was used to assess initial feasibility and acceptability. Detailed exit interviews were conducted to elicit participant feedback, which helped inform the development of a substantially upgraded in-house IVR system. This refined IVR system was then subjected to a sequential explanatory mixed methods evaluation. Participating rural county coordinators and research staff (N=10) tested the usability of the IVR system features for 2 weeks and then completed the System Usability Scale and qualitative semistructured interviews.

Results: The study sample comprised mostly African American people, women, rural county coordinators, and research staff (N=10). Participants rated the IVR system with a mean score of 81 (SD 5) on the System Usability Scale, implying *excellent* usability. In total, 5 overarching themes emerged from the qualitative interviews: likes or dislikes of the intervention, barriers to or facilitators of PA, technical difficulties, quality of calls, and suggestions for intervention improvement. Message framing on step feedback, call completion incentives, and incremental goal-setting challenges were areas identified for improvement. The positive areas highlighted in the interviews included the personalized call schedules, flexibility to call in or receive a call, ability to make up for missed calls, narration, and PA tips.

Conclusions: The usability testing and feedback received from the rural county coordinators and research staff helped inform a final round of refinement to the IVR system before use in a large randomized controlled trial. This study stresses the importance of usability testing of all digital health interventions and the benefits it can offer to the intervention.

(JMIR Hum Factors 2022;9(1):e29494) doi:10.2196/29494

KEYWORDS

interactive voice response systems; usability; exercise; physical activity; rural health; telehealth

Introduction

Background

Automated telephone-based intervention strategies may be key to overcoming the numerous barriers to physical activity (PA) promotion and cancer control in the Alabama Black Belt, a rural region named for its rich soil but whose population is at increased risk for sedentary lifestyles and related cancer disparities [1]. Low literacy, poverty, lack of transportation, cultural preferences, and distance from PA facilities often impede access to PA information and resources in this region [2]. Interactive voice response (IVR) systems allow users to interact by pressing keys on the telephone keypad and can be effective in targeting behavior change [3]. The recent National Health Interview Survey estimates that only 0.7% of the population in the United States is phoneless [4], thereby demonstrating the potential for a wider reach of IVR interventions. Moreover, IVRs do not require clinic visits, high literacy, or access to costly technology [5,6].

In response, we have developed an IVR-delivered PA intervention that is currently being tested in a large randomized controlled trial (RCT) in 6 rural Black Belt counties of Alabama. This paper describes the process that led us to the design of the IVR system and the results of the usability testing that was conducted before the commencement of the RCT.

As with any intervention, particularly digital health interventions, examination of the usability of the developed intervention before the actual deployment of the intervention is vital [7]. With IVR systems featuring only voice-based output and keypad-based input, a seamless user experience can indeed be tricky [8-10]. IVR-based intervention systems can pose more challenges than simple IVR data collection systems, as IVR-based intervention systems need to focus on achieving minimal information navigation time, while featuring maximal information relevance and capacity [8].

Objectives

Our proposed study aims to target rural Black Belt counties of Alabama that are marked by low literacy and education levels [1,2]. Although there is a body of work focused on the usability of IVR systems [8-10], there is limited research on the usability of IVR systems for rural settings and underserved populations. This limited body of literature has used surveys and interviews to evaluate the usability of IVR systems. This study seeks to fill this gap in the literature by reporting our development methodology, system features, and explanatory sequential mixed methods design to assess the usability of the IVR system. Our hypotheses are that most participants in this usability study will rate the usability of the IVR system favorably and provide useful suggestions for further improvements during interviews.

Methods

Parent Study Overview

The parent study (R01CA233550) is an ongoing RCT (N=240) comparing a Deep South IVR-Supported Active Lifestyle (DIAL) intervention with a waitlist control among underactive adults residing in 6 rural Alabama counties [11]. On the basis of the social cognitive theory (SCT) [12], this study extends an IVR-supported PA intervention that targets key SCT constructs (self-regulation, self-efficacy, enjoyment, outcome expectations, and social support) through IVR counseling calls. The participants are provided pedometers (Accusplit AX2790MV) and Fitbit activity monitors (model: Inspire) to record daily steps and receive progress feedback via the IVR PA-tracking and goal-setting calls.

The number of calls in a week tapers as participants progress through the intervention (from daily calls in months 0-3 to twice per week in months 4-6 and weekly in months 7-12), and the content of the calls vary based on specific days of the intervention.

Iterative IVR System Design

Piloting a Beta Version of IVR

A previous pilot study (R03CA177538) tested a beta version of this IVR system with a convenience sample (N=63) [13-15]. Findings from this trial supported the feasibility and acceptability of the approach and helped further refine the technology and theory-driven intervention components in preparation for extension to rural populations. More specifically, the findings yielded the need for IVR-initiated calls as opposed to only participant-initiated calls, specific targeting of unchanged SCT constructs and incorporating multi-level strategies (incremental goal-setting and county coordinator support) for increased support, accountability, and sustainability [14].

The IVR system used in this pilot study was a commercial IVR system and posed several limitations. First, all voice clips were prerecorded by voice narrators and uploaded. Second, the system only worked by participants calling into the system and did not offer a way for the system to initiate calls. The commercial system also posed limitations in terms of dynamic tailored questions that used earlier responses to frame newer questions as the call progressed.

Upgrading and Refining the IVR System

In response to this pilot study feedback, we developed a completely homegrown IVR system for the parent RCT using

more up-to-date technology. This new system was hosted on a Linux server, powered by an Apache web server, programmed using the Laravel framework (a hypertext preprocessor–based rapid development framework), data stored using a MySQL database, and connected to Twilio for telephony.

Although the system was being developed, we conducted focus groups with multiple stakeholder groups (rural county coordinators and research staff from the University of Alabama at Birmingham [UAB] O'Neal Comprehensive Cancer Center Community Outreach and Engagement Office). For the focus groups, we generated 3 sample voice clips of intervention messages using Amazon Polly, a text-to-speech engine, and presented the 3 sample voice clips to our stakeholders. Amazon Polly is capable of close to human-like voices, which resulted in stakeholders preferring Amazon Polly voices over prerecorded human voices. This choice of Amazon Polly voices also allows for the use of different tones and genders for the voices during the calls and avoids the extensive time and financial costs associated with rerecording message libraries with human narrators every time an edit is made to the content.

The focus group participants also provided feedback on incoming versus outgoing calls, the preferred procedure to handle missed calls, and other support strategies. More specifically, they felt that their community members and potential participants would appreciate the flexibility and convenience of bidirectional calls and the option to fill in missed call data at later dates. The need to be able to change phone numbers and allow incoming calls from new (unregistered) numbers was stressed. Support strategies, such as brief counseling sessions during in-person data collection and offering Fitbit devices were also suggested.

The system development was conducted in an agile fashion, with regular demonstrations to the rural country coordinators from the UAB O'Neal Comprehensive Cancer Center Community Outreach and Engagement Office. Their feedback regarding the speed of the voice clips, pauses between sentences, pauses in sentences, length of the phone call, reading level of the language used in the calls, and logical flow of the content resulted in numerous edits. The system included a participant call completion incentive mechanism that awarded the participants a minimum of US $0.25 for each call completed. However, the incentive amount became US $0.50 when the participant completed 7 preceding calls, with the incentive falling back to US $0.25 when a call was not completed.

The development phase concluded with the core project staff (DP, MT, ST, and VR) pilot-testing the revised system to identify and fix any problems. Some examples of the problems identified and fixed include system expecting responses within 5 seconds, incorrect feedback messages, and outgoing calls not being placed as scheduled. After this, a formal usability test was conducted as detailed in the following section. Finally, the system went through another round of iterative refinements based on the findings from the usability testing. The details of the resultant system are presented in the *Results* section.

Usability Testing

Study Design

This study incorporated an explanatory sequential mixed methods design to assess the usability of an IVR phone counseling system that will be extended to physically inactive residents in 6 rural Alabama counties (Hale, Choctaw, Greene, Marengo, Dallas, and Sumter). Demographics were assessed at baseline. System usability and semistructured interviews were conducted at the 2-week follow-up.

Participants

The sample for usability testing comprised 10 rural county coordinators and research staff affiliated with the UAB O'Neal Comprehensive Cancer Center Community Outreach and Engagement Office who would later serve a critical role in recruitment, assessment, and intervention delivery for the RCT study but had yet to be exposed to the newly developed IVR system.

Procedures

Each participant completed a one-on-one orientation via Zoom with the DIAL program manager or principal investigator. During the session, the participants were given an overview of the usability study protocols and the IVR system, completed an initial IVR call with the research team, and asked questions.

Following orientation, the participants began wearing a study-provided pedometer or an approved personal activity monitor (ie, Fitbit or Apple Watch) and receiving daily IVR calls from DIAL for 2 weeks. The participants received all 3 types of IVR calls: PA-tracking, goal-setting, and counseling calls. For tracking calls, the participants answered PA questions (reported pedometer use, steps per day, and any moderate-intensity PA in the past 24 hours) and received PA tips and feedback. Tracking calls lasted approximately 1 minute per call. During the counseling calls, the participants answered PA questions and additional questions covering PA self-efficacy, enjoyment, outcome expectations, and social support. Moreover, they received tailored feedback on these psychosocial variables based on their individual responses to these questions. Counseling calls lasted approximately 10 minutes per call. Goal-setting calls allowed the participants to set their own step goal or increase their current step goal by 500 steps for the upcoming week. Goal-setting calls lasted approximately 5 minutes per call. In the 2-week period, the participants received 1 call per day, with a total of 1 goal-setting call, 1 counseling call, and 12 PA-tracking calls.

Quantitative Measures

Demographics

Participant demographics, including age, gender, educational attainment, race and ethnicity, household income, employment, marital status, and number of children living at home were assessed at baseline.

Survey Items

At follow-up, the participants completed the System Usability Survey (10 items) on the web via Qualtrics XM (Qualtrics) combined with 4 more project-specific items. All 14 items were

aimed at assessing how the participants felt about the phone counseling system after using it for 2 weeks. The participants responded to the statement—*Please select the answer that best expresses how you feel about each statement after using the phone counseling system over the past 2 weeks*—for items such as *I think I would like to use this phone counseling system frequently*, *I thought the phone counseling system was easy to use*, *I felt very confident using the phone counseling system*, and *I needed to learn a lot of things before I could get going with this phone counseling system*. The 4 project-specific items were worded as *How likely are you to recommend this system to others?*, *Did you receive your calls at the scheduled time?*, *What gender was the voice on your calls?*, and *Did you use the study-provided pedometer to track your steps?*

Quantitative Analysis

All quantitative data collected during this study were descriptively analyzed. Microsoft Excel was used for all the quantitative analyses.

Qualitative Methodology

After 2 weeks of receiving calls and completing the quantitative survey, all 10 rural county coordinators and research staff participated in one-on-one, semistructured interviews conducted via Zoom regarding their experiences with the calls and how usability could be improved before implementing the IVR for the RCT. The semistructured interview guide was developed by coauthors (DP and SN) and included questions regarding motivation to exercise, likes and dislikes of the calls, specific call features that could motivate or demotivate individuals, technical aspects of the IVR call, and suggestions for improvement. To ensure consistency, all interviews were conducted in July 2020 by 1 member of the study team (SN) with expertise and experience in qualitative interviewing. SN is not involved in any aspect of the broader RCT or technology design and development and was engaged to serve as a neutral evaluator for the purpose of this usability evaluation.

Qualitative Analysis

All interviews were audio-recorded and transcribed verbatim by a professional transcription service. Thematic analysis [16] was conducted using NVivo 13 (QSR International) [17]. Investigator triangulation methodology was conducted [18] by a 2-member analysis team (DP and SN) with experience in qualitative methodology in social science disciplines (clinical psychology and medical sociology) who independently reviewed transcripts through line-by-line coding. After the initial categories and themes were generated in a cyclical, iterative process, the full research team refined the existing categories, themes, and subthemes. Discrepancies, although infrequent, were addressed with the research team.

Results

Participant Characteristics

Participant characteristics are shown in Table 1. The total sample included 10 participants with an average age of 48.7 (SD 18.6) years. The sample comprised largely female (8/10, 80%), Black (8/10, 80%), and non-Hispanic or Latino (9/10, 90%) participants, and had no children living at home (9/10, 90%). Most reported completing college (6/10, 60%) and either full-time or part-time employment at the time of usability testing. Half of the sample (5/10, 50%) reported <US $50,000 annual household income, and only 30% (3/10) of the participants reported never being married.

Table 1. Demographic characteristics of the usability testing participants (N=10).

Demographics	Values
Age (years), mean (SD)	48.7 (18.6)
Sex, n (%)	
Female	8 (80)
Race, n (%)	
Black or African American	8 (80)
Non-Hispanic or Latino	1 (10)
Other	1 (10)
Education, n (%)	
Some college	4 (40)
College graduate	3 (30)
Postgraduate work	3 (30)
Household annual income (US $), n (%)	
<50,000	5 (50)
≥50,000	5 (50)
Marital status, n (%)	
Single (never married)	3 (30)
Married	4 (40)
Divorced	3 (30)
Children living at home, n (%)	
None	9 (90)
≥1	1 (10)

Quantitative Results

The usability testing survey that was conducted after 2 weeks of IVR system use yielded positive results (Table 2). All participants (10/10, 100%) agreed that the IVR system was easy to use without the need for technical assistance or extensive learning, and most (7/10, 70%) would recommend the IVR system to others. The participants were confident in using IVR (8/10, 80%), and 70% (7/10) would like to use IVR frequently. Very few participants found the IVR system cumbersome (2/10, 20%) or confusing (3/10, 30%), and only 10% (1/10) of the participants found the IVR system to be unnecessarily complex. In terms of functionality, 70% (7/10) of the participants agreed that the various functions of the IVR system were well-integrated. The participants (7/10, 70%) reported receiving their calls at the scheduled time, and 90% (9/10) reported a female voice on their calls. Only 40% (4/10) of the participants reported wearing the study pedometer; however, of the 60% (6/10) who did not wear the study pedometer, 50% (3/6) used an Apple Watch and 50% (3/6) used a Fitbit Inspire. To numerically interpret the usability of the system, the standardized System Usability Scale scoring procedure was used [19]. This resulted in an average score of 81 (SD 5). Previous research indicates that a System Usability Scale score of >68 can be considered as above-average usability. This score of 81 translates to an e*xcellent* usability rating [20].

Table 2. Usability testing survey results (N=10).

Statement and answers[a]	Participants, n (%)
I think I would like to use this phone counseling system frequently.	
Somewhat agree	7 (70)
Neither agree nor disagree	3 (30)
I found the phone counseling system unnecessarily complex.	
Strongly disagree	7 (70)
Somewhat disagree	2 (20)
Somewhat agree	1 (10)
I thought the phone counseling system was easy to use.	
Strongly agree	5 (50)
Somewhat agree	5 (50)
I think that I would need the support of a technical person to be able to use this system.	
Strongly disagree	7 (70)
Somewhat disagree	3 (30)
I found the various functions in this phone counseling system were well-integrated.	
Strongly agree	3 (30)
Somewhat agree	4 (40)
Somewhat disagree	3 (30)
I thought there was too much inconsistency in this phone counseling system.	
Strongly disagree	4 (40)
Somewhat disagree	3 (30)
Neither agree nor disagree	2 (20)
Somewhat agree	1 (10)
I would imagine that most people would learn to use this phone counseling system very quickly.	
Strongly agree	4 (40)
Somewhat agree	6 (60)
I found the phone counseling system very cumbersome to use.	
Strongly disagree	4 (40)
Somewhat disagree	4 (40)
Neither agree nor disagree	1 (10)
Somewhat agree	1 (10)
I felt very confident using the phone counseling system.	
Strongly agree	6 (60)
Somewhat agree	2 (20)
Neither agree nor disagree	1 (10)
Somewhat disagree	1 (10)
I needed to learn a lot of things before I could get going with this phone counseling system.	
Strongly disagree	5 (50)
Somewhat disagree	5 (50)
How likely are you to recommend this system to others? (Scale of 0-10)	
2	2 (20)
6	1 (10)
8	4 (40)

Statement and answers[a]	Participants, n (%)
9	2 (20)
10	1 (10)
Did you receive your calls at the scheduled time?	
Yes	7 (70)
No	3 (30)
What gender was the voice on your calls?	
Female	9 (90)
Both male and female	1 (10)
Did you use the study-provided pedometer to track your steps?	
Yes	4 (40)
No	6 (60)
Apple Watch	3 (30)
Fitbit Inspire	3 (30)

[a]Please select the answer that best expresses how you feel about each statement after using the phone counseling system over the past 2 weeks.

Qualitative Results

A total of 5 overarching themes emerged: (1) likes or dislikes of the intervention, (2) barriers to or facilitators of PA, (3) technical difficulties, (4) quality of the calls, and (5) suggestions for improvement of the intervention.

Likes and Dislikes About the IVR Intervention

When asked what they liked about the DIAL intervention, several participants stated that the phone calls motivated them to exercise and kept them accountable:

> *I wasn't as active, but after I went through the calls, I became more active and aware, and I was becoming used to the calls, and I was looking forward to the calls, and I was looking forward to the motivational tips.*

> *I think because it held me accountable. The accountability to hear what I had accomplished and what I not accomplished, that adds extra value because it almost puts a mirror in front of your face, and says, "Look." Sometimes it's very difficult to look at that mirror, and say, "This is what I have or have not done."*

The participants also appreciated the flexibility of the (new) bidirectional call format:

> *One thing I did like was that, for instance, if I did not make my call. I had the opportunity to call back. That was good.*

> *You have the different options and different times of calling, I think that's good for the people that's busy. So if they miss the call, they can call back, or the system will call them back, but if they need to change their time of the call, then they able to do that.*

Finally, the participants looked forward to the PA tips at the end of the call:

> *I think motivating tips at the end, they were good. I knew them already, but I listened to them. So I think that they were good for people that's just starting out with their health journey.*

Regarding dislikes, the participants expressed concerns that specific step feedback messages were *negative* and *stern*. For example, when <10,000 steps per day were reported in the PA-tracking call, the participants received the following feedback:

> *Thanks for reporting your steps. You did not meet the DIAL study step goal of 10,000 steps per day yet, but you are on your way. Keep making small increases until you get there.*

The participants had strong reactions to this feedback and compared it to *a slap in the face*:

> *It would be a little discouraging to hear that every day, "You didn't meet your 10,000 steps goal. You did not meet the goal. You did not meet the goal."*

The incentives for IVR call completion were another dislike, particularly for rural county coordinators:

> *I don't think that that 25 cents is helpful for motivating people to continue to get the call.*

> *What's with the incentive? That's kind of really, make you feel a little worthless.*

Other participants were more open to the idea:

> *Anything that's an incentive that would give people the extra motivation to want to do it, I think it's a good idea...it's not much, but it gives you that sense of, "I made it. I got a quarter, I got 50 cents." It's not much, but it's that knowing that something is in [it at] the end for you.*

Facilitators of and Barriers to Participating in PA or the IVR Intervention

Chronic disease prevention and management was an important motivator for participation in PA or the IVR intervention:

> *I think depending on where people are in their lives, being physically active might be motivated by so saying, "Hey, this disease process can be kept at bay or managed or maybe even prevented if you exercise."*

Social support was also key to encouraging PA initiation and maintenance, especially once the DIAL intervention ended:

> *Yes, I do think that they will start or to continue to exercise if they have a friend or a buddy to walk with or whatever. I think that that is important to have someone to exercise with.*

> *I think that the interpersonal aspect of it will be really important. Although it's not a person, when that connection and accountability with the phone system is removed, I think it will be really important to have that from another source, and hopefully other participants or family members of the participants.*

As for barriers to engaging in PA and completing the IVR calls, the participants stressed the lack of time and competing interests:

> *There are many, many days where I don't want to do any physical activity. I would say actually most days. It's not because it's tedious. It's because I have so many things to do, and I keep thinking, "Wow. I got to spend that hour doing this."*

Technical Difficulties

The participants described experiencing some initial technical difficulties with the IVR calls, such as receiving calls at incorrect times or with *system error* messages. The programming decision to skip calls on holidays also seemed to cause some confusion and was changed as a result:

> *During the 4th of July holidays, I didn't receive any calls at all that weekend.*

Finally, the participants learned to take their time entering the responses during the IVR calls:

> *If you trying to speed it up and hurry up, you know you going to press two, nuh-uh. it's going to hit you with an error.*

Quality of the Calls

The participants generally indicated that the *quality of the call is good*. In fact, rural county coordinators had previously given the Amazon Polly narration a favorable review at a focus group. During usability testing, several participants distinguished the female voice options as *less monotone*, *robotish* than certain *boring* male voice options. The pace of the calls received mixed reviews; *It really was a good pace* for some and *a bit too fast* for others:

> *There were times where it felt like it was moving a bit too fast, especially when there were multiple options or the question or the prompt was read or said, and then the answers were said immediately after. I don't know, sometimes it was rushed through, it felt like.*

Suggestions

The participants stated that they preferred to have written *user-friendly* instruction materials that could be used during the calls:

> *I think you can give them a little prompt card. Like some of the prompts. Because it's just different every time, but just a small, little introduction of what to expect.*

The participants also suggested that having printed materials of the survey readily available to community participants would be beneficial and crucial for capturing accurate data:

> *I think it would help to have some sort of printout. Just a scale that says, "This is what one means, and this is what 100 means." Because again, and maybe I just was doing too many things sometimes. For me to remember what that scale was. If I had a call that was coming through, of course I would not click over, but that's a distraction for me, again. So if I'm in the middle of a question, I'm like, "Oh my gosh. Am I supposed to pick 1 or 100." I think a printout scale or something in front of me probably would have reminded me, because again, some days I'm putting one, some days I'm putting 99, and that's not what I meant. But there was no way for me to go back and erase my answer, to my knowledge. I don't know if I missed that in the training, but I just thought, "Oops." But that was just my short term memory, and knowing I needed to complete the call.*

In addition to the advantage of obtaining accurate data, the participants stated that having printed materials would also mitigate noncompliance from frustration:

> *Anything that we could get to assist would always be helpful. So, if we could come up with something, some type of visual aid for the older generation, then that will be great. I'm sure everybody can work a phone, but you want to make sure that they're not getting confused. Because once they get confused, confusion causes discouraged sometimes. So, you don't want to get them confused. So, yeah, if we could come up with some type of handout that would be great.*

The participants suggested that step goals should be modest and community participants should be reminded to gradually increase their steps during the study:

> *Take baby steps. I think if you want to see somebody make it to that 10,000 goal marker, try to start small, like see where they're at and see what is an average for the participant and then work your way up from there. And ultimately, it may be just too hard for somebody's daily schedule to meet that 10,000 goal step without making significant changes to their daily routines. But I think if you take baby steps, then for a good bit of the participants, you may not got to do it for everybody, but at least a certain group of the*

> *participants, and you're able to get some changes, like an increase in daily steps from them, they're all meeting the goals, then I think that could be a good motivating factor.*
>
> *It did tell me, try to add a 500 steps for the next time, but I feel like it should be more personalized like, "Okay, you got 2000, tomorrow let's try to get 2,500." And then when they call the next day, if they have the 2,500 be like, "Great, you met the goal. Do you think you could add another 500?" I feel like that's how it should be. I don't feel like the bar should right off the top be 10,000 because that's a lot for some people.*

The participants also provided solutions for the previously mentioned issues with incentives for call completion (eg, substituting nonmonetary incentives and a point system):

> *I can tell you that getting 25 cents for each phone call was not motivating at all. Yeah. I don't mean to be blunt, but it wasn't. It wasn't motivating. How much I earned at the end of each phone call, it just didn't motivate me. Now, if I earned points for each phone call, and I could redeem those points in some sort of physical activity, online store.*
>
> *I mean, the value may be still $3.75, but with let's just say 375 points may be like a gift. It may be a pedometer. It may be a little lunch tote, or it may be something else. People can use a tote. They can use a pedometer. They can use even a cup or a mug, or a water bottle if they're exercising. But what can you do with $3.75?*

Finally, the participants discussed the tips provided to increase the number of steps and suggested that they be personalized to the Deep South rural community:

> *For instance, tell them like get up during commercial break and walk around your coffee table twice. That's feasible. It's within reach and it doesn't take a lot of effort to go out. Because think about these people that don't have parks nearby. We're telling them to go to the nearest park. Well, there is no nearest park.*
>
> *Tell them, okay, well walk around your house two times or walk to the mailbox twice or for instance what's something else. Go three mailboxes down and come back. Something that people can be like, "Oh yeah, I can do that. I never thought to do that."*

Discussion

Summary

Innovative IVR systems hold the potential to overcome barriers to achieving the recommended levels of PA in the rural Black Belt region of Alabama [3]. However, no previous research has examined IVR systems in rural contexts to increase PA levels. We developed an IVR system in an iterative manner based on feedback from earlier pilot studies, focus groups, and the current usability testing with key stakeholders (both community members and local county coordinators with UAB O'Neal Comprehensive Cancer Center Community Outreach and Engagement Office). The resultant system was characterized by high usability and is currently being tested for efficacy in an RCT.

Principal Findings and Resultant IVR System

The IVR system received a numerical usability score of 81—equating to an e*xcellent* usability score. The sequential explanatory mixed methods design we adopted helped us identify several opportunities for improvement through the qualitative interviews. After usability testing (qualitative interviews), we implemented several improvements into the system. First, we modified our messaging when the participants failed to reach their goals to sound less negative or stern. We implemented graceful handling of wrong key presses by participants; instead of informing them that they had pressed a wrong key, we reworded to say that the system could not understand. We implemented a detailed orientation session procedure in which the IVR system was oriented and printed materials were made available. Instead of directly pushing the participants toward 10,000 steps, following participant feedback, we implemented incremental goals of 250 steps per week. We reworded our reward system to use the word *points* instead of *cents* to emphasize the gamification of IVR adherence versus financial transactions. Finally, we also added several more PA tips as suggested by our interview participants.

As a means to further test the system before the commencement of the RCT, the core group of researchers working on this study met to discuss whether further formal usability testing was required. As most reported issues pertained to wording or content, it was decided that no further usability testing was needed. However, the core group of researchers were listed as pilot users of the IVR system. These pilot users were scheduled approximately 2-3 weeks ahead of the actual RCT participants. These researchers actively tested the system daily and reported to the development team on any issue found. This enabled the development team to aggressively address the issues before any RCT study participants encountered them. Some example issues identified and fixed using this approach included problems when the participants moved from one phase (daily calls) to another (biweekly calls) and nonavailability of new PA strategies to suggest to participants. Our *2 weeks ahead approach* enabled us to resolve these issues before any real participants encountered them while avoiding lapses in time that would delay the project.

Our final product is a comprehensive IVR system with cutting-edge capabilities such as streamlined calls, smart dropped calls handling, and assignable voice gender. Future research should examine the added value of such features and their impact on this promising technology.

Final IVR System Design

The feedback during, before, and after the usability testing was used to iteratively refine the IVR system. The resultant system, which is now being used in the RCT, is described as follows:

1. The system is designed to handle complex call schedules involving different types of calls during different phases of the intervention and the randomization group.

2. The system can receive incoming calls and smartly place outgoing calls only if the participant has a pending incomplete call.
3. To protect the privacy of the participants, they are identified using their phone number and a personal identification number (PIN). When the participants use their registered phone, only a PIN is required. When the participants use a phone other than their registered phone, both the registered phone number and PIN are required. This achieves a balance between user experience and security.
4. New participants are registered on a web portal by a study manager who retrieves the unique PIN for the user. The study manager is then able to print the PIN and other instructional materials in an educational binder for the participants.
5. A comprehensive missed call policy has been implemented, with the system retrying the call after 30 minutes. Again, if there is no response, the call is marked as incomplete and can be completed the next day.
6. A smart dropped-call policy has also been implemented, wherein if a participant drops midway through a call and the user connects again within a preset time limit, the participant is able to continue from the last question they answered.
7. One of the most significant aspects affecting the usability of IVR systems is the information navigation time [7]. We have essentially eliminated the navigation time by streaming the content for calls in multiple ways. First, calls are not placed unless there is a pending survey to be completed. Second, when users have multiple surveys due, the system combines all the surveys and offers them in a sequence. Finally, if the participant has any pending surveys as a result of missed calls in the previous 2 days, the system offers the missed surveys in sequence.
8. Many IVR systems require a significant amount of time because of the confirmation messages, such as "You pressed 6, press 1 if this is right or press 2 to change." These confirmation messages are necessary as it is easy for a participant to accidentally mistype a number; however, these confirmation messages almost double the call time. To overcome this, during the orientation session, we educate the users on pressing * anytime during the call to edit the last response.
9. The participants can call the IVR system anytime and change their preferred call receiving time.
10. To maximize information relevance [7], the system is programmed to be able to look up the participant's previous step goals, PA self-efficacy, enjoyment, social support, and outcome expectancies and use those values as a part of the conversation—thereby leading to high relevance.
11. To maximize information capacity [7], a bank of PA-increasing strategies has been created, with new strategies being revealed on a weekly basis. Similarly, a bank of greeting messages has also been made available. Through these mechanisms, despite the daily calls during the first 3 months, the users would find a variety of content being delivered.
12. At the end of each call, the system announces the reward points earned by the participant, which can be redeemed for actual monetary incentives. Before the usability testing, we directly referred to points as cents. However, we learned from usability testing that the participants felt that 25 cents per call made them feel that their time was worthless. Thus, we reworded our call content to award points rather than cents.
13. The gender of the voice narration in the call can be set to match the gender of the participant, to the opposite gender, or to be random.

Strengths and Limitations

This study had a few limitations. First, usability testing was conducted during the COVID-19 pandemic; thus, for participant safety, all surveys and interviews were conducted remotely, and participation was limited to community health advisors and staff. Although these community health advisors live and work in the same rural counties of the Black Belt region of Alabama as the future participants, it is possible that they do not accurately represent the demographics of the participants (eg, education levels) who would participate in the RCT study. In addition, our demographics includes predominantly female and non-Hispanic or Latino populations.

However, this opportunity allowed rural county coordinators to gain familiarity and comfort with the inner workings of the IVR system before spearheading its dissemination among their own communities. Thus, they will be more prepared to orient participants to the IVR system and field their questions. Moreover, playing such a key role in the development and refinement of this technology likely enhanced the sense of buy-in and ownership among these key stakeholders and gatekeepers to the community and substantially improved the final product.

Conclusions

This study demonstrated that the developed IVR system is usable and has the potential to increase the levels of PA. Study findings provided insight into the participants' preferred language, narration tones, rewards, and variety of messaging. These insights can be valuable for future studies that seek to develop IVR-based interventions.

Acknowledgments

This study was funded in part by the National Cancer Institute (R01CA233550) and the National Heart, Lung, and Blood Institute training program (T32HL105349).

Authors' Contributions

MT led the development of the interactive voice response system and was thus not involved in the usability testing of the system. DP and SN, who were not part of the interactive voice response system development team, conducted the usability tests and interviews.

Conflicts of Interest

None declared.

References

1. Gyawu R, Quansah JE, Fall S, Gichuhi PN, Bovell-Benjamin AC. Community food environment measures in the Alabama Black Belt: implications for cancer risk reduction. Prev Med Rep 2015 Aug 22;2:689-698 [FREE Full text] [doi: 10.1016/j.pmedr.2015.08.015] [Medline: 26844138]
2. Pekmezi D, Marcus B, Meneses K, Baskin ML, Ard JD, Martin MY, et al. Developing an intervention to address physical activity barriers for African-American women in the deep south (USA). Womens Health (Lond) 2013 May;9(3):301-312 [FREE Full text] [doi: 10.2217/whe.13.20] [Medline: 23638785]
3. Tsoli S, Sutton S, Kassavou A. Interactive voice response interventions targeting behaviour change: a systematic literature review with meta-analysis and meta-regression. BMJ Open 2018 Feb 24;8(2):e018974 [FREE Full text] [doi: 10.1136/bmjopen-2017-018974] [Medline: 29478016]
4. Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January – June 2021. National Center for Health Statistics. URL: https://www.marketscreener.com/news/latest/Wireless-Substitution-Early-Release-of-Estimates-from-the-National-Health-Interview-Survey-January--37035906/ [accessed 2021-11-27]
5. Kassavou A, Sutton S. Automated telecommunication interventions to promote adherence to cardio-metabolic medications: meta-analysis of effectiveness and meta-regression of behaviour change techniques. Health Psychol Rev 2018 Mar;12(1):25-42. [doi: 10.1080/17437199.2017.1365617] [Medline: 28805162]
6. Posadzki P, Mastellos N, Ryan R, Gunn L, Felix L, Pappas Y, et al. Automated telephone communication systems for preventive healthcare and management of long-term conditions. Cochrane Database Syst Rev 2016 Dec 14;12:CD009921 [FREE Full text] [doi: 10.1002/14651858.CD009921.pub2] [Medline: 27960229]
7. Lewis J. Usability testing. In: Handbook of Human Factors and Ergonomics, Third Edition. Hoboken, New Jersey: John Wiley & Sons; 2006.
8. Asthana S, Singh P, Singh A. Exploring the usability of interactive voice response system's design. In: Proceedings of the 3rd ACM Symposium on Computing for Development. 2013 Presented at: Proceedings of the 3rd ACM Symposium on Computing for Development; Jan 11 - 12, 2013; Bangalore India. [doi: 10.1145/2442882.2442924]
9. Cilliers L, Flowerday S. Factors that influence the usability of a participatory IVR crowdsourcing system in a smart city. South African Comput J 2017 Dec 08;29(3). [doi: 10.18489/sacj.v29i3.422]
10. Lerer A, Ward M, Amarasinghe S. Evaluation of IVR data collection UIs for untrained rural users. In: Proceedings of the First ACM Symposium on Computing for Development. 2010 Presented at: Proceedings of the first ACM symposium on computing for development; Dec 17 - 18, 2010; London United Kingdom. [doi: 10.1145/1926180.1926183]
11. Brown NI, Powell MA, Baskin M, Oster R, Demark-Wahnefried W, Hardy C, et al. Design and rationale for the deep south interactive voice response system-supported active lifestyle study: protocol for a randomized controlled trial. JMIR Res Protoc 2021 May 25;10(5):e29245 [FREE Full text] [doi: 10.2196/29245] [Medline: 34032575]
12. Bandura A. Human agency in social cognitive theory. Am Psychol 1989 Sep;44(9):1175-1184. [doi: 10.1037/0003-066x.44.9.1175] [Medline: 2782727]
13. Ainsworth M. Effects of Interactive Voice Response (IVR) counseling on physical activity benefits and barriers. Health Behav Policy Rev 2020 Oct;7(5):407-415. [doi: 10.14485/hbpr.7.5.3]
14. Pekmezi D, Ainsworth C, Holly T, Williams V, Joseph R, Wang K, et al. Physical activity and related psychosocial outcomes from a pilot randomized trial of an interactive voice response system-supported intervention in the deep south. Health Educ Behav 2018 Dec;45(6):957-966 [FREE Full text] [doi: 10.1177/1090198118775492] [Medline: 29884069]
15. Pekmezi D, Ainsworth C, Holly T, Williams V, Benitez T, Wang K, et al. Rationale, design, and baseline findings from a pilot randomized trial of an IVR-Supported physical activity intervention for cancer prevention in the Deep South: the DIAL study. Contemp Clin Trials Commun 2017 Dec;8:218-226 [FREE Full text] [doi: 10.1016/j.conctc.2017.10.008] [Medline: 29503878]
16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
17. Zamawe F. The implication of using NVivo software in qualitative data analysis: evidence-based reflections. Malawi Med J 2015 Mar;27(1):13-15 [FREE Full text] [doi: 10.4314/mmj.v27i1.4] [Medline: 26137192]
18. Thurmond VA. The point of triangulation. J Nurs Scholarsh 2001;33(3):253-258. [doi: 10.1111/j.1547-5069.2001.00253.x] [Medline: 11552552]

19. Brooke J. SUS: a "quick and dirty" usability scale. In: Usability Evaluation In Industry. Boca Raton, Florida: CRC Press; 1996.
20. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 May 1;4(3):114-123. [doi: 10.5555/2835587.2835589]

Abbreviations

DIAL: Deep South IVR-Supported Active Lifestyle
IVR: interactive voice response
PA: physical activity
PIN: personal identification number
RCT: randomized controlled trial
SCT: social cognitive theory
UAB: University of Alabama at Birmingham

Edited by A Kushniruk; submitted 08.04.21; peer-reviewed by H Belli, F Ghezelbash, F Yu; comments to author 04.07.21; revised version received 29.08.21; accepted 11.09.21; published 04.01.22.

Please cite as:
Thirumalai M, Brown N, Niranjan S, Townsend S, Powell MA, Neal W, Schleicher E, Raparla V, Oster R, Demark-Wahnefried W, Pekmezi D
An Interactive Voice Response System to Increase Physical Activity and Prevent Cancer in the Rural Alabama Black Belt: Design and Usability Study
JMIR Hum Factors 2022;9(1):e29494
URL: https://humanfactors.jmir.org/2022/1/e29494
doi:10.2196/29494
PMID:34982714

Original Paper

Multisensory Home-Monitoring in Individuals With Stable Chronic Obstructive Pulmonary Disease and Asthma: Usability Study of the CAir-Desk

Dario Kohlbrenner[1,2], PhD; Christian F Clarenbach[1], MD; Adam Ivankay[3], MSc; Lukas Zimmerli[3], PhD; Christoph S Gross[4], PhD; Manuel Kuhn[1,2], MSc; Thomas Brunschwiler[3], PhD

[1]Department of Pulmonology, University Hospital Zurich, Zurich, Switzerland

[2]Faculty of Medicine, University of Zurich, Zurich, Switzerland

[3]IBM Research Europe - Zurich, Zurich, Switzerland

[4]Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

Corresponding Author:
Dario Kohlbrenner, PhD
Department of Pulmonology
University Hospital Zurich
Raemistrasse 100
Zurich, 8091
Switzerland
Phone: 41 43 253 08 42
Email: dario.kohlbrenner@usz.ch

Abstract

Background: Research integrating multisensory home-monitoring in respiratory disease is scarce. Therefore, we created a novel multisensory home-monitoring device tailored for long-term respiratory disease management (named the CAir-Desk). We hypothesize that recent technological accomplishments can be integrated into a multisensory participant-driven platform. We also believe that this platform could improve chronic disease management and be accessible to large groups at an acceptable cost.

Objective: This study aimed to report on user adherence and acceptance as well as system functionality of the CAir-Desk in a sample of participants with stable chronic obstructive pulmonary disease (COPD) or asthma.

Methods: We conducted an observational usability study. Participants took part in 4 weeks of home-monitoring with the CAir-Desk. The CAir-Desk recorded data from all participants on symptom burden, physical activity, spirometry, and environmental air quality; data on sputum production, and nocturnal cough were only recorded for participants who experienced symptoms. After the study period, participants reported on their perceptions of the usability of the monitoring device through a purpose-designed questionnaire. We used descriptive statistics and visualizations to display results.

Results: Ten participants, 5 with COPD and 5 with asthma took part in this study. They completed symptom burden questionnaires on a median of 96% (25th percentile 14%, 75th percentile 96%), spirometry recordings on 55% (20%, 94%), wrist-worn physical activity recordings on 100% (97%, 100%), arm-worn physical activity recordings on 45% (13%, 63%), nocturnal cough recordings on 34% (9%, 54%), sputum recordings on 5% (3%, 12%), and environmental air quality recordings on 100% (99%, 100%) of the study days. The participants indicated that the measurements consumed a median of 13 (10, 15) min daily, and that they preferred the wrist-worn physical activity monitor to the arm-worn physical activity monitor.

Conclusions: The CAir-Desk showed favorable technical performance and was well-accepted by our sample of participants with stable COPD and asthma. The obtained insights were used in a redesign of the CAir-Desk, which is currently applied in a randomized controlled trial including an interventional program.

(JMIR Hum Factors 2022;9(1):e31448) doi:10.2196/31448

KEYWORDS

home monitoring; digital health; respiratory disease; usability; feasibility; adherence; disease management; chronic disease; patient monitoring

Introduction

Respiratory disease has a huge impact on global health. Taken together, both acute and chronic respiratory disease rank second among all-cause mortality [1]. At the same time, chronic respiratory disorders carry substantial disease burden. Chronic obstructive pulmonary disease (COPD), a major chronic respiratory condition, poses a substantial risk of disability (the ninth most frequent cause of disability worldwide) [1]. Global networks of experts engage in improving diagnostic approaches and researching the most effective treatment modalities for chronic respiratory disease [2,3]. Current evidence indicates multimodal diagnostic and treatment approaches to be the method of choice.

Technological advances substantially impact health care. Home-monitoring and self-reporting allow more frequent measurements of individuals' symptoms and behaviors, supported by seamless data transfer, storage, and analysis in the cloud, enabling on-demand overviews of subjects' health states and trajectories. However, cutting-edge technology seems scarcely implemented into clinical practice and clinical research. The most widely used approach is "tele-healthcare," in which part of the management is carried out through telephone calls [4-7]. This is an interesting option in mobility-compromised populations or for minor health issues. However, this approach is still staff-intense and provides a unidimensional focus on solely patient-reported outcomes (PROMs). Furthermore, appointments, and accordingly the time points when the PROMs are measured, are most commonly dictated by the schedule of the health care professionals and not by the time point at which participants experience symptoms, attacks, barriers, or insecurities. We hypothesize that recent technological accomplishments can be integrated into a multisensory participant-driven platform, which involves both PROMs and disease markers incorporating multiple measures from a variety of sensors. We also believe that this platform could improve chronic disease management and be accessible to large groups at an acceptable cost.

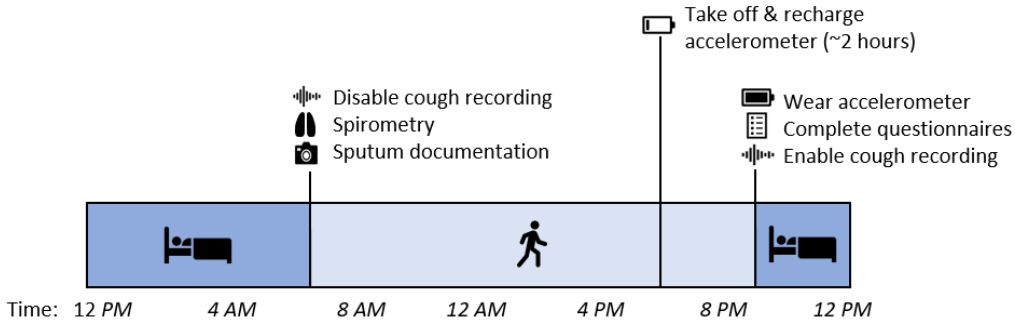
Research integrating multisensory home-monitoring in respiratory disease is scarce [8]. The available investigations were most commonly unidimensional, short-term, and small-scale. To overcome these shortcomings, we developed a novel multisensory home-monitoring device tailored for long-term respiratory disease management (named the "CAir-Desk"), which is optimized for user experience, health workflow, and outcomes. In this study, we report on system functionality of the novel CAir-Desk and feasibility, as well as user adherence and acceptance in a sample of participants with stable COPD or asthma.

Methods

Study Design

We conducted an observational usability study. Participants conducted 4 weeks of disease home-monitoring with the CAir-Desk, without interventions or modifications to their established treatment regimen. The participants were instructed to place the CAir-Desk in their bedroom and take daily measurements in accordance with a schedule (Figure 1). After the study period, participants reported their perceptions on the usability of the CAir-Desk with a purpose-designed questionnaire (Multimedia Appendix 1). This study did not fall within the scope of the Human Research Act (HRA) [9] and did not require authorization from the ethics committee. The Ethics Committee of the Canton of Zurich confirmed this in the BASEC Request (2018-00180).

Figure 1. Daily measurement schedule during the study period.



Study Participants

We used convenience sampling for this study. Ten participants with stable COPD or asthma, attending outpatient secondary care at the Department of Pulmonology, University Hospital Zurich, Switzerland, consented to take part in the study.

The CAir-Desk

The CAir-Desk (Figure 2), is a novel, custom-built disease home-monitoring system [10]. It combines multiple sensors in a compact format with a single power plug for device charging. All components of the CAir-Desk are conformité européenne (CE)-certified.

Figure 2. The CAir-Desk setup for the usability study as a collection of sensors. Items indicated here are as follows: (a) smartphone, (b1) arm-worn accelerometer, (b2) wrist-worn accelerometer, (c) spirometer, (d) sputum collector and smartphone camera, (e) environmental air quality monitor, and (f) nocturnal cough monitor.

Smartphone

The smartphone (Galaxy A320, 2017, Samsung Group)—object *a* in Figure 2—contained purpose-designed apps for user interaction and data visualization. All sensors are accessible via the smartphone. Further, all sensor data are transferred to the cloud storage through the smartphone by the Global System for Mobile Communications (GSM) network.

Physical Activity

Physical activity was measured using multisensory triaxial accelerometry (Charge 3, Fitbit Inc, and Everion, Biovotion). We decided to include 2 physical activity monitors for this usability study to conclude on the participants' preference. The Charge 3 is a wrist-worn device (item *b2* in Figure 2) with a small display that provides real-time information on step count. The Everion does not include a display and is worn on the upper arm, (item *b1* in Figure 2). It only provides information when synchronized with the CAir-Desk.

Symptom Burden

Symptom burden in the participants diagnosed with COPD was assessed using the COPD Assessment Test [11]. For the participants diagnosed with asthma, a purpose-designed questionnaire was used (Multimedia Appendix 2). The questionnaires were sent out through the patient-provider communication channel app (docdok.health). The smartphone displayed push notifications when the daily questionnaire was due, and participants answered the questions directly in the app.

Nocturnal Cough

Nocturnal cough recordings were collected using the smartphone microphone (item *f* in Figure 2). The participants had to manually enable recording every night before going to bed and turn it off in the mornings. A purpose-designed algorithm isolated and extracted the cough count from the background noise. The CAir-Desk app displayed cough count in running monitoring sessions in near real-time (10-second delay) [12]. This was a voluntary measurement; we encouraged only participants with a self-reported cough to use the sensor.

Spirometry

Daily spirometry recordings were obtained with a portable spirometry device (Air Next Spirometer, NuvoAir)—item *c* in Figure 2—which connected with the smartphone via Bluetooth. The values obtained were forced expiratory volume in 1 second (FEV_1) and forced vital capacity (FVC). All participants were trained to produce reproducible maneuvers that comply with published guidelines [13].

Sputum Monitoring

This was a voluntary measurement; we encouraged only those participants with self-reported sputum production to submit a photograph of a sample to the system. This was accomplished using the built-in smartphone camera and purpose-designed sputum collectors (item *d* in Figure 2).

Environmental Air Quality

The CAir-Desk continuously recorded environmental air quality (Foobot, Airboxlab)—item *e* in Figure 2. The values obtained were temperature, humidity, particulate matters, and volatile organic compounds. No interaction with the CAir-Desk was needed from the participants to enable data-recording.

Cloud and Backend Setup

Data transfer to and data storage on the cloud were performed with encryption, as required by data protection regulations for sensitive personal information.

Further details about technical, cloud, and backend solutions can be found elsewhere [10].

Device Usage and Adherence Thresholds

To identify days of adherence, we defined individual thresholds for each sensor in accordance with the study protocol. An overview of these thresholds for each sensor, together with details on recording modalities, is provided in Table 1.

Table 1. Overview of the sensor and adherence thresholds.

Sensor	Relevant modalities	Adherence day threshold
Mandatory[a]		
Physical activity, wrist-worn	Step count	Steps ≥ 100
Physical activity, arm-worn	Step count	Recording ≥ 1 hour
Symptom burden	CAT[b] [11] or Asthma questionnaire	Completed questionnaire
Spirometry	Forced expiratory volume in 1 second, forced vital capacity	≥3 valid exhalations
Environmental air quality	Volatile organic compounds, temperature	Recording ≥ 1 hour
Voluntary[c]		
Nocturnal cough	Audio (.wav) file count	Recording ≥ 1 audio file
Sputum	Photo count	Recording ≥ 1 photo

[a]Measurements were requested from all participants.

[b]CAT: chronic obstructive pulmonary disease assessment test.

[c]Measurements were requested from participants experiencing specific symptoms.

Statistical Analysis

This small-scale usability study did not allow assumptions on normally distributed data, which we confirmed visually using quantile-quantile plots. The day of instruction and the day of return of the CAir-Desk were both excluded from the analysis, only considering full study days. In case of technical issues, for example, when a sensor was not recording for the whole or part of the study period, these data were excluded from adherence-reporting.

We extracted and prepared the data for statistical analysis with Python 3.7.10 (The Python Software Foundation, 2021). We analyzed the data statistically with R 4.0.3 (R Core Team, 2021, R Foundation for Statistical Computing).

Results

All results are presented using descriptive statistics in the format median (25th percentile, 75th percentile) unless otherwise stated.

Ten participants, 5 with COPD (61 [59, 62] years; median FEV_1 predicted=69% [60%, 78%]) and 5 with asthma (55 [55, 55] years; median FEV_1 pred=62% [62%, 69%]) were included in this study (detailed characteristics in Table 2) and used the CAir-Desk for 27.5 (27, 28) days, excluding the on- and off-boarding days. All participants completed the predetermined study period and did not experience any adverse events.

Table 2. Study participant characteristics stratified in accordance with their diagnosis.

Characteristics	Participants with chronic obstructive pulmonary disease (n=5)	Participants with asthma (n=5)
Age (years), median (25th percentile, 75th percentile)	61 (59, 62)	55 (55, 55)
Sex (female/male), n (%)	3/2 (60/40)	1/4 (25/75)
Forced expiratory volume in 1 second (% predicted), median (25th percentile, 75th percentile)	69 (60, 78)	62 (62, 69)
Smoking status (yes/no), n (%)	1/4 (25/75)	0/5 (0/100)

Participant Adherence and Technical Considerations

Questionnaires

In total, participants completed 96% (14%, 96%) of the daily questionnaires. Participants with COPD completed 96% (0%, 96%) and those with asthma completed 96% (22%, 96%) of the daily questionnaires. Owing to a technical issue, one participant with COPD did not receive the questionnaires and was excluded from the analysis of adherence data. Figure 3A provides individual adherence rates throughout the study period; Figure 3A indicates participants who could not adhere owing to technical difficulties. Except for 2 days, adherence rates with respect to the study days remained constantly high throughout the study period (Figure 4A).

Figure 3. Individual daily adherence for symptom burden questionnaires (A) and spirometry (B). All participants are shown (ie, including those with nonadherence due to technical difficulties). COPD: chronic obstructive pulmonary disease.

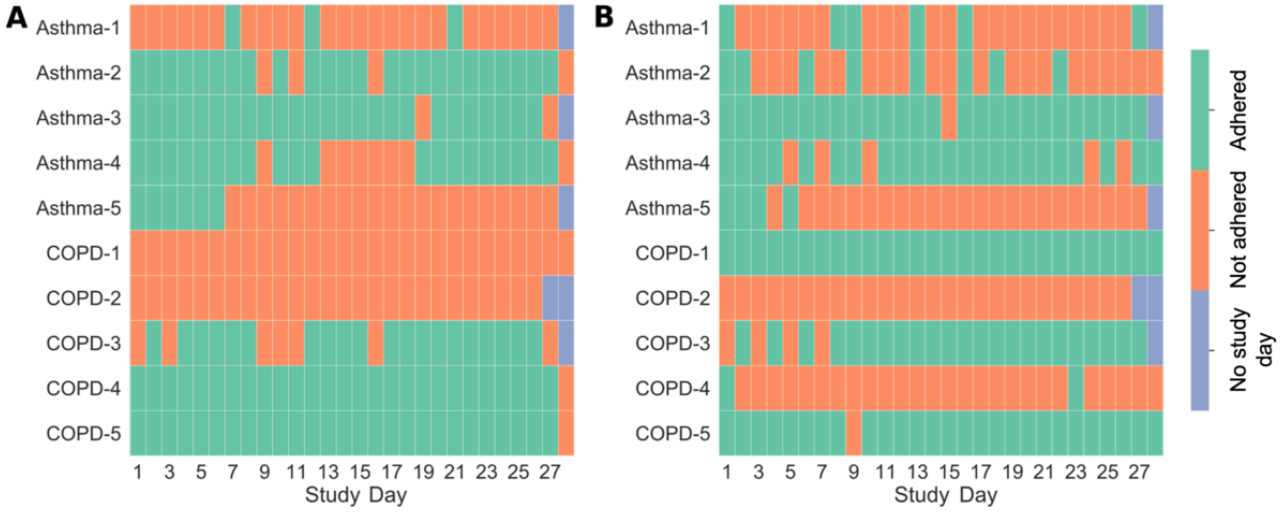
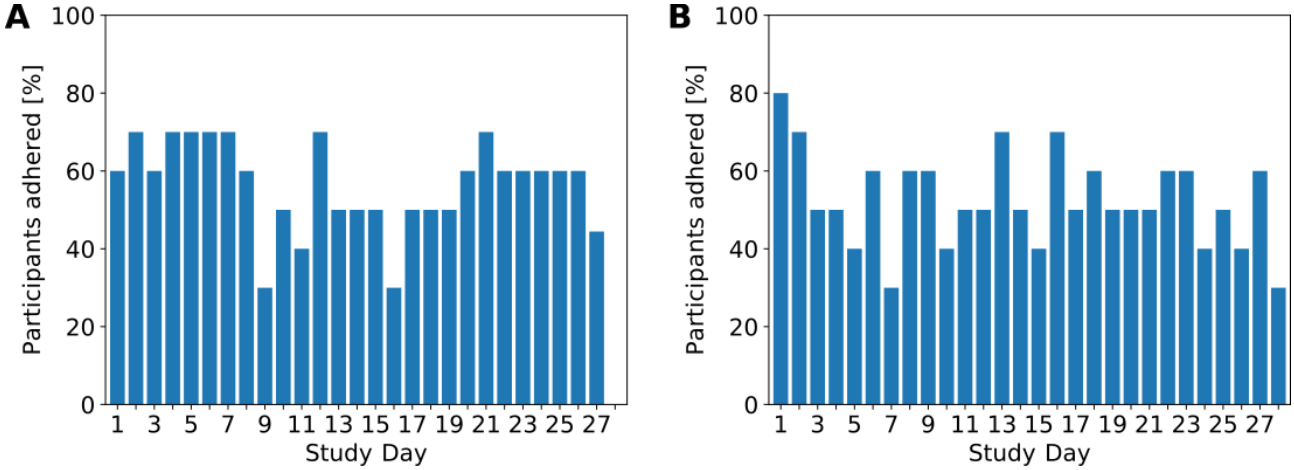


Figure 4. Daily adherence (ratio of participants who adhered to the total number of study participants) for symptom burden questionnaires (A) and spirometry (B). All participants are represented (ie, including those with nonadherence due to technical difficulties).



Spirometry

Participants completed valid spirometry recordings (ie, 3 or more attempts) on 55% (20%, 94%) of the study days. Participants with COPD completed valid spirometry recordings on 96% (48%, 96%) and participants with asthma completed recordings on 29% (22%, 82%) of the study days. For 2 participants with COPD, the data upload partially failed during the study period. These participants were excluded from the analysis of adherence data. Figure 3B shows individual adherence rates throughout the study period; Figure 3B includes participants who could not adhere owing to technical difficulties. There was variability in daily adherence rates throughout the study period with more consistency toward the end of the study (Figure 4B).

Physical Activity

Participants wore the wrist-worn accelerometer on 100% (97%, 100%) of the study days. Participants with COPD wore the wrist-worn accelerometer on 100% (100%, 100%) and those with asthma wore it on 100% (96%, 100%) of the study days.

Participants wore the arm-worn accelerometer on 45% (13%, 63%) of the study days. Participants with COPD wore the arm-worn accelerometer on 61% (30%, 80%) and those with asthma wore it on 30% (14%, 63%) of the study days. For 2 participants with COPD, the arm-worn accelerometer broke during daily use and could not be used further; these participants were excluded from the analysis of adherence data.

Nocturnal Cough Recording

Participants enabled nocturnal cough recording on 34% (9%, 54%) of the study nights. Participants with COPD enabled nocturnal cough–recording on 46% (23%, 59%) and those with asthma enabled it on 22% (11%, 50%) of the study nights. For 2 participants with COPD, the data upload did not proceed as expected throughout part of the study period; these participants were excluded from the analysis of adherence data.

Sputum Monitoring

Participants took photographs of their sputum on 5% (3%, 12%) of the study days. Participants with COPD took photographs of their sputum on 4% (2%, 16%) and those with asthma on 7% (4%, 11%) of the study days. For 2 participants with COPD, the data upload did not proceed as expected throughout part of the study period; these participants were excluded from the analysis of adherence data.

Environmental Air Quality Monitoring

Participants measured environmental air quality on 100% (99%, 100%) of the study days. Participants with COPD measured environmental air quality on 100% (100%, 100%) and those with asthma on 100% (96%, 100%) of the study days. For 2 participants with COPD, the data upload did not proceed as expected during part of the study period; these participants were excluded from the analysis of adherence data. One participant with asthma turned off the data transfer connection for the sensor by accident for a short time, which resulted in 15 days of nonrecording. We retained this participant in the analysis.

Adherence data on the questionnaires, spirometry recordings, physical activity, nocturnal cough recording, sputum monitoring, and environmental monitoring are presented on individual and summary levels in Table 3 and are visualized in Figure 5.

Table 3. Sensor adherence data on a participant[a] and summary[b] level.

Participant number	Diagnosis	Mandatory[c]					Voluntary[d]	
		Questionnaire	Spirometry	Physical activity, wrist-worn	Physical activity, arm-worn	Air quality	Nocturnal cough	Sputum
1	Asthma	11	22	78	63	96	11	7
2	Asthma	96	29	100	14	50	68	0
3	Asthma	96	93	100	30	100	4	11
4	Asthma	96	82	100	64	100	50	4
5	Asthma	22	15	96	7	100	22	15
Summary of participants with asthma	N/A[e]	96 (22, 96)	29 (22, 82)	100 (96, 100)	30 (14, 63)	100 (96, 100)	22 (11, 50)	7 (4, 11)
6	COPD[f]	0	100	100	61	100	46	4
7	COPD	N/A	0	96	100	100	0	0
8	COPD	96	N/A	100	N/A	N/A	N/A	N/A
9	COPD	96	N/A	100	N/A	N/A	N/A	N/A
10	COPD	96	96	100	0	100	71	29
Summary of participants with COPD	N/A	96 (0, 96)	96 (48, 98)	100 (100, 100)	61 (30, 80)	100 (100, 100)	46 (23, 59)	4 (2, 16)
Summary total	N/A	96 (14, 96)	55 (20, 94)	100 (97, 100)	45 (13, 63)	100 (99, 100)	34 (9, 54)	5 (3, 12)

[a]Individual data are percentage values of the total number of study days.

[b]Summary data are median (25th percentile, 75th percentile) percentage values of the total number of study days.

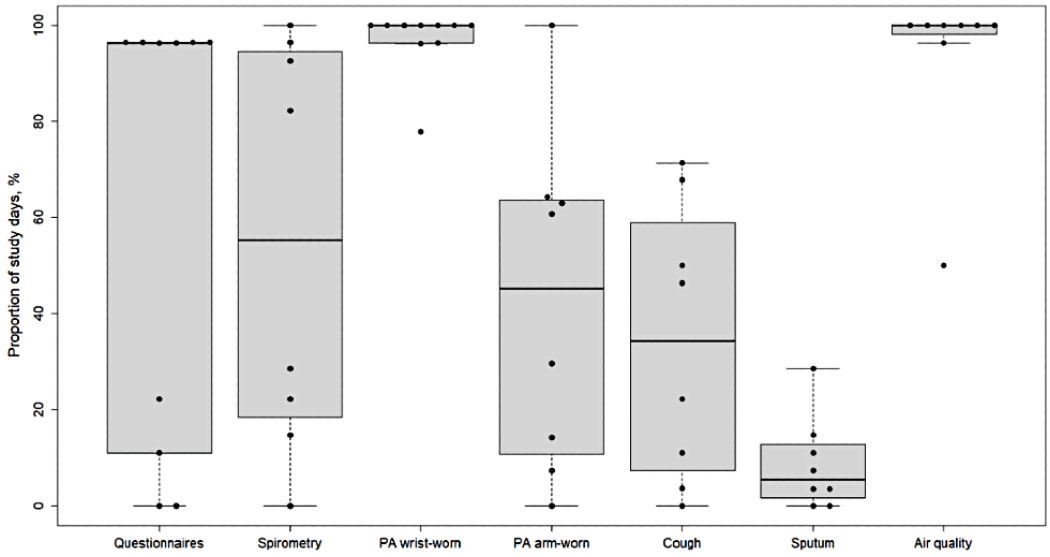
[c]Measurements were obtained from all participants.

[d]Measurements were only requested from participants who experienced specific symptoms.

[e]N/A: not applicable.

[f]COPD: Chronic obstructive pulmonary disease.

Figure 5. Adherence stratified for sensors. Data are the percentages of study days. PA, physical activity.



Participant Satisfaction

Regarding the CAir-Desk, in general, 6 (60%) participants indicated that they would be willing to use the device further as is, and 1 (10%) participant indicated that he would not. Furthermore, 2 (20%) participants indicated a neutral response, and 1 participant (10%) did not answer the question. Overall, participants reported to have taken a median of 13 (10, 15) minutes per day manipulating the device, which 7 (70%) participants considered an acceptable amount of time, 2 (20%) participants considered too much time, and 1 (10%) participant indicated a neutral response.

Questionnaires

Concerning further use, 4 (40%) participants indicated that they would be willing to continue daily reporting, and 4 (40%) participants indicated that they would not. Two (20%) participants indicated a neutral response.

Regarding user-friendliness, 4 (40%) participants considered the questionnaires user-friendly, and 6 (60%) did not.

Spirometry

Regarding further use, 4 (40%) participants indicated that they would be willing to continue daily recordings, and 3 (30%) participants indicated that they would not. Two (20%) participants indicated a neutral response, and 1 (10%) participant did not use the spirometry device.

Regarding user-friendliness, 7 (70%) participants considered the spirometry device user-friendly, and 3 (30%) did not.

Physical Activity

Regarding further use of the wrist-worn accelerometer, 9 (90%) participants indicated that they would be willing to continue wearing the device, and 1 (10%) participant indicated that he would not.

Regarding user-friendliness, 9 (90%) participants considered the wrist-worn accelerometer user-friendly and 1 (10%) did not.

Regarding further use of the arm-worn accelerometer, 5 (50%) participants indicated that they would be willing to continue wearing the device, and 4 (40%) indicated that they would not. Furthermore, 1 (10%) participant indicated a neutral response.

Regarding user-friendliness, 5 (50%) participants considered the arm-worn accelerometer user-friendly and 5 (50%) did not.

Nocturnal Cough Recording

Regarding further use, 4 (40%) participants indicated that they would be willing to continue recordings, and 4 (40%) indicated that they would not. Furthermore, 2 (20%) participants indicated a neutral response.

Regarding user-friendliness, 4 (40%) participants considered the nocturnal cough recording user-friendly and 6 (60%) did not.

Sputum Monitoring

Regarding further use, 3 (30%) participants indicated that they would be willing to continue taking photographs, and 2 (20%) participants indicated that they would not. Furthermore, 2 (20%) participants indicated a neutral response, and 2 (20%) did not use this function owing to no sputum production. One (10%) participant did not answer this question.

Regarding user-friendliness, 3 (30%) participants considered sputum monitoring user-friendly and 7 (70%) did not.

Environmental Air Quality Monitoring

Regarding further use, 8 (80%) participants indicated that they would be willing to continue air quality monitoring, and 2 (20%) indicated a neutral response.

Regarding user-friendliness, 5 (50%) participants considered environmental air quality monitoring user-friendly and 5 (50%) did not.

Participant satisfaction data on the questionnaires, spirometry, physical activity, nocturnal cough recording, sputum monitoring, and environmental monitoring are displayed in Figure 6. Adherence to questionnaires, spirometry, and physical activity monitoring stratified in accordance with the participants' rating on user-friendliness (Figure 7).

Figure 6. Participant satisfaction data stratified for sensors. (A) “Would you be willing to use [insert sensor] further?”. (B) “Do you consider [insert sensor] as user-friendly?” Data are n (on the x-axis) and % (on the bars). PA: physical activity.

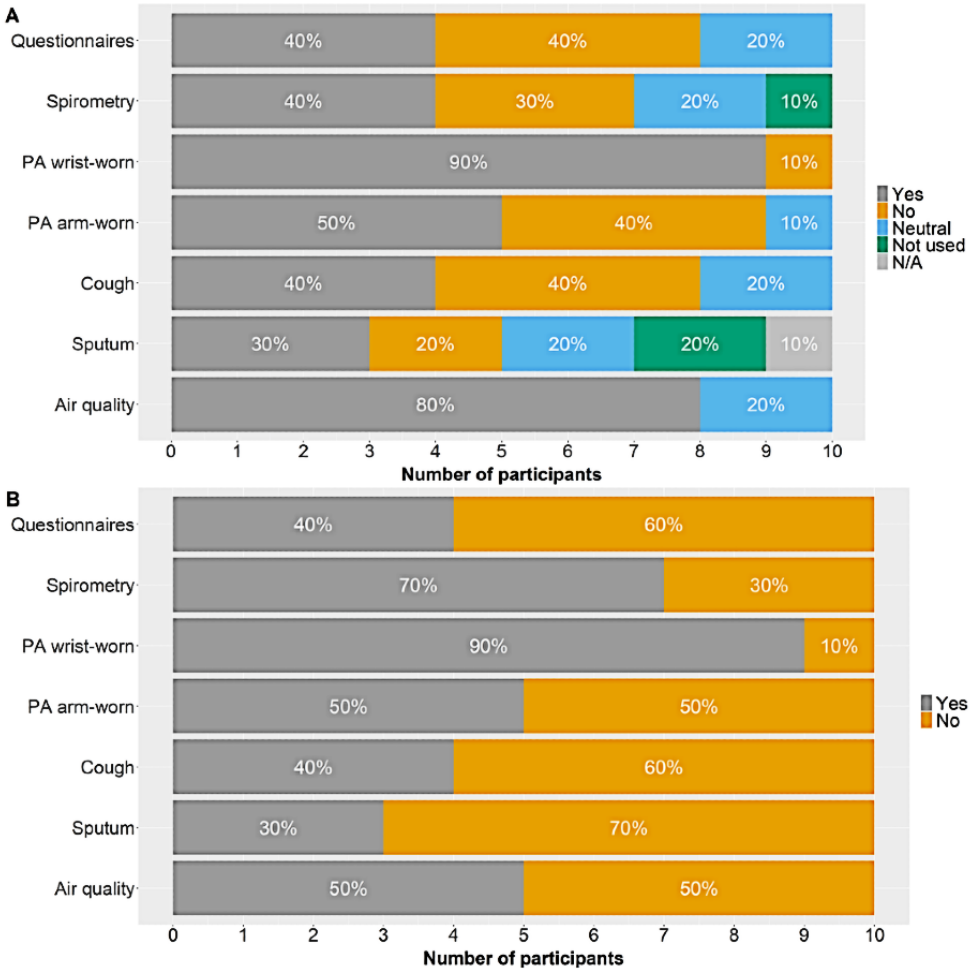
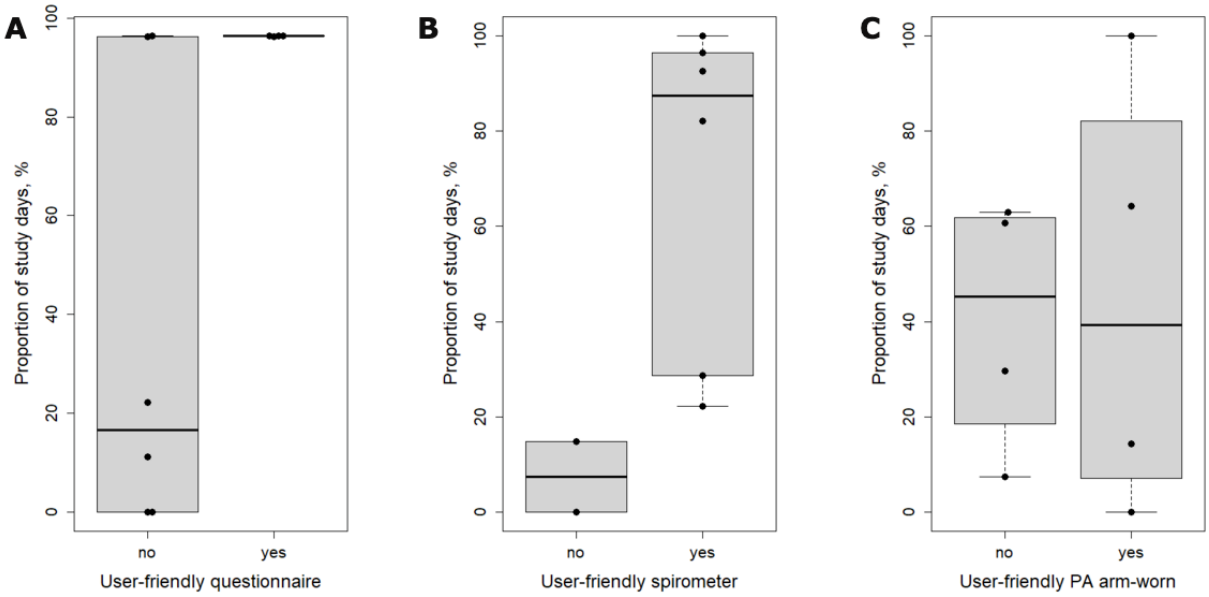


Figure 7. Adherence stratified according to the user-friendliness of sensors. Questionnaires (A), spirometry (B), arm-worn accelerometer (C). Data are the percentages of study days. PA: physical activity.



Discussion

Principal Findings

We report on the first clinical application of the CAir-Desk. In this usability study, the CAir-Desk performed well technically and was well-accepted by our sample of participants with stable COPD and those with asthma.

General Considerations

In general, participants indicated that they would be willing to use the CAir-Desk further. Since this was a purely observational study, we conclude that participants, although in stable phases of COPD or asthma, are interested in disease monitoring. The time consumption that the participants reported was lower than expected and, for the vast majority (80%), did not exceed their personal tolerance. We consider this a positive finding, since we performed the usability study in preparation for a subsequent randomized controlled trial (RCT) on the effectiveness of delivering "Living well with COPD" by a chatbot through the CAir-Desk, in combination with the home-monitoring [10,14]. When receiving an intervention through the CAir-Desk, time spent manipulating the device will certainly increase. However, we expect the amount of time to still be acceptable given our present findings.

General adherence to the intense measurement schedule was considerably high. Inconsistency in adherence rates between individuals was observed in spirometry and nocturnal cough measurements, which was due to the voluntary nature of these measurements.

Considering that this was the first clinical application of the CAir-Desk, technical performance and robustness was satisfactory. We faced problems when transferring data to the cloud with 2 participants, which were resolved remotely (however, leading to incomplete adherence data). Regarding hardware robustness, participants reported 2 broken arm-worn accelerometers during the study time. These had to be replaced and reconnected to the CAir-Desk by study staff, which led to an unplanned additional study visit and incomplete adherence data.

In response to the in-field testing, the CAir-Desk was updated for the subsequent RCT by including a script, which opens the relevant sensor apps daily, to guarantee data synchronization even when apps are closed unintentionally. Furthermore, we decided to use the wrist-worn accelerometer in any further clinical application of the CAir-Desk.

Detailed Considerations

Regarding the daily questionnaires that the participants were asked to answer, we found high adherence, although the willingness for further use and user-friendliness did not receive consistently high ratings. One participant was not able to answer the questionnaires because the font was too small for him to read. For this usability study, we recruited individuals with stable COPD or asthma. Therefore, we hypothesize that daily administration of questionnaires might be too frequent since only minor daily variation in symptoms and general health status are expected. However, we would still recommend daily questionnaires in samples that are not in a stable disease phase, receive a new intervention, or are newly diagnosed. We recommend using a short questionnaire, keeping time consumption low and adherence high. Finally, large fonts or zoom functionality on the questionnaires is crucial to allow completion for individuals with compromised vision. Since this is a feasibility study investigating a very small sample, some caution on the interpretation is needed. This applies especially to the questionnaires in the case of the participants with COPD, since conclusive data are available from only three participants.

Regarding spirometry, we observed high interindividual variability. While one participant with COPD never took a test, another participant adhered strictly to the schedule. Overall, the adherence with respect to study days was almost constant across the study period (Figure 4B). Unfortunately, upload problems in 2 participants prevented us from drawing stronger conclusions on adherence patterns. However, we consider acceptance to daily spirometry tests as high, with only 2 participants reporting that they would not be willing to perform further tests. Similar to the questionnaires, we hypothesize that individuals with unstable disease or recent diagnosis would be more interested in daily measurements.

Regarding physical activity monitoring, we integrated two different devices for this usability study. We were interested if the participants preferred the arm-worn device or the wrist-worn device. The main differences between the devices are visibility to others, live access to measurements through a built-in display, and battery life. Participants clearly preferred the wrist-worn device (with built-in display and 5-day battery life instead of 1 day for the arm-worn device), which was also reflected in adherence rates. Only one participant reported the wrist-worn accelerometer not to be user-friendly, because the display turned on during sudden movements at night and caused awakenings. From a technical point of view, both devices can record heart rate. However, validity is low and may not yet serve medical purposes [15].

Regarding nocturnal cough recording, adherence was highly variable because only participants experiencing cough were encouraged to take measurements. User-friendliness was considered low. Participants had to turn on the cough counter manually every time they went to bed and turn it off when they got up. The majority of the participants considered this not user-friendly and we therefore updated the application following the study and added the feature for automated recording during individually tailorable time periods.

Regarding sputum monitoring, only the few participants who were experiencing sputum production were encouraged to take measurements. Again, we consider the option to submit a photograph of sputum samples as valuable in individuals with unstable disease. However, the CAir-Desk is nonportable; therefore, information on sputum production when individuals are not at home is missed. This has to be considered when populations are younger, and more active (eg, participants with cystic fibrosis).

Regarding environmental air quality monitoring, adherence was high, and participants indicated interest in further use. The device recorded automatically, and we only faced some days

of nonrecording for 1 participant who accidentally turned off the Wi-Fi hotspot connection for several days. Unfortunately, the app displaying the results was only available in English. This fact led our participants, all German speakers, to low user-friendliness ratings.

Our findings suggest that missing data in similar studies to ours may be categorized into four categories: (1) missing owing to nonadherence (eg, did not like to use the sensor), (2) missing owing to technical issue (eg, failing upload or broken device), (3) missing owing to other issue (eg, allergy to worn device or impaired vision), (4) missing owing to absence of symptoms. Furthermore, it may be assumed that the participants who did not rate a sensor as user-friendly show lower adherence to that sensor. Our data show that this is not necessarily the case (Figure 7). While there was a clear trend to lower adherence in daily spirometry measurements among participants who did not consider the sensor user-friendly, a mixed response was observed in daily questionnaires, and an overall low adherence was observed in the arm-worn accelerometry measurements.

Considering the findings of this usability study, we modified the CAir-Desk, the corresponding sensors, apps, and the data upload. The modified version of the CAir-Desk is now used in an RCT delivering the multimodal intervention "Living well with COPD" through a chatbot [10].

Conclusions

Our investigation indicates that multisensory disease monitoring using the CAir-Desk is feasible, and well accepted. Additionally, time consumption was surprisingly low and suggests that intervention delivery through the CAir-Desk can be performed with a reasonable time budget. Our results add knowledge on important points to consider when designing multisensory setups for individuals with chronic respiratory disease. First, sensors and apps should use automation as often as possible. Second, participants are interested in seeing their measurements and performance. Therefore, apps should grant easy interaction and provide an intuitive overview on the recorded information in the participant's preferred language. Third, a rigid schedule with daily measurements is feasible and accepted by the participants. However, tailoring in accordance with the disease stage seems important to increase adherence in longer studies. Last, technical issues were rare and manageable.

We believe that the combination of home-based multisensory measurements and PROMs might be a game changer in chronic disease management. The frequent and accurate data obtained with the CAir-Desk might provide novel insights into early markers for a decline in health status or emergencies. We consider the combination of home-based measurements and information delivery combined with regular in-person care as the future of multimodal chronic disease management.

Acknowledgments

We thank the participants who took part in this study and helped us to improve the CAir-Desk for the subsequent trial. This study was funded by Innosuisse (project 29844.1).

Authors' Contributions

CFC and TB conceived and designed the study. AI, LZ, and TB designed and implemented the CAir-Desk. DK and MK collected the data. DK and TB extracted the data and performed the statistical analysis. DK and CFC recruited the participants. DK drafted the manuscript. CFC, AI, LZ, CSG, MK, and TB critically revised and approved the manuscript.

Conflicts of Interest

CFC has received remuneration for lectures and talks, development of educational material, participation in advisory boards, and travel and accommodation or meeting reimbursements from Roche, Novartis, Boehringer, GSK, Astra Zeneca, Sanofi, Vifor, OM Pharma, and Mundipharma within the last 36 months, all outside the submitted work. All other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Usability questionnaire.
[DOCX File , 158 KB - humanfactors_v9i1e31448_app1.docx]

Multimedia Appendix 2
Asthma symptom burden questionnaire.
[DOCX File , 117 KB - humanfactors_v9i1e31448_app2.docx]

References

1. Disease burden, 2000–2019. World Health Organization. URL: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates-leading-causes-of-dalys [accessed 2022-02-01]
2. Global Strategy for Asthma Management and Prevention (2020 update). Global Initiative for Asthma. 2020. URL: https://ginasthma.org/wp-content/uploads/2020/06/GINA-2020-report_20_06_04-1-wms.pdf [accessed 2022-02-01]

3. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: 2021 Report. Global Initiative for Chronic Obstructive Pulmonary Disease, Inc. URL: https://goldcopd.org/wp-content/uploads/2020/11/GOLD-REPORT-2021-v1.1-25Nov20_WMV.pdf [accessed 2022-02-03]
4. Rassouli F, Pfister M, Widmer S, Baty F, Burger B, Brutsche MH. Telehealthcare for Chronic Obstructive Pulmonary Disease in Switzerland Is Feasible and Appreciated by Patients. Respiration 2016;92(2):107-113 [FREE Full text] [doi: 10.1159/000448377] [Medline: 27553807]
5. Rassouli F, Germann A, Baty F, Kohler M, Stolz D, Thurnheer R, et al. Telehealth mitigates COPD disease progression compared to standard of care: a randomized controlled crossover trial. J Intern Med 2021 Mar;289(3):404-410 [FREE Full text] [doi: 10.1111/joim.13230] [Medline: 33428219]
6. Anthony Jnr B. Implications of telehealth and digital care solutions during COVID-19 pandemic: a qualitative literature review. Inform Health Soc Care 2021 Mar 02;46(1):68-83. [doi: 10.1080/17538157.2020.1839467] [Medline: 33251894]
7. Schutte-Rodin S. Telehealth, Telemedicine, and Obstructive Sleep Apnea. Sleep Med Clin 2020 Sep;15(3):359-375. [doi: 10.1016/j.jsmc.2020.05.003] [Medline: 32762969]
8. Janjua S, Banchoff E, Threapleton CJ, Prigmore S, Fletcher J, Disler RT. Digital interventions for the management of chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2021 Apr 19;4:CD013246. [doi: 10.1002/14651858.CD013246.pub2] [Medline: 33871065]
9. Federal Act on Research involving Human Beings. The Federal Assembly of the Swiss Confederation. 2011 Sep 30. URL: https://www.admin.ch/opc/de/classified-compilation/20061313/index.html [accessed 2022-02-01]
10. Gross C, Kohlbrenner D, Clarenbach CF, Ivankay A, Brunschwiler T, Nordmann Y, et al. A Telemonitoring and Hybrid Virtual Coaching Solution "CAir" for Patients with Chronic Obstructive Pulmonary Disease: Protocol for a Randomized Controlled Trial. JMIR Res Protoc 2020 Oct 22;9(10):e20412 [FREE Full text] [doi: 10.2196/20412] [Medline: 33090112]
11. Jones PW, Harding G, Berry P, Wiklund I, Chen W, Kline Leidy N. Development and first validation of the COPD Assessment Test. Eur Respir J 2009 Sep;34(3):648-654 [FREE Full text] [doi: 10.1183/09031936.00102509] [Medline: 19720809]
12. Vhaduri S, Kessel T, Ko B, Wood D, Wang S, Brunschwiler T. Nocturnal Cough and Snore Detection in Noisy Environments Using Smartphone-Microphones. 2019 Presented at: IEEE International Conference on Healthcare Informatics (ICHI); June 10-13, 2019; Xi'an. [doi: 10.1109/ichi.2019.8904563]
13. Graham BL, Steenbruggen I, Miller MR, Barjaktarevic IZ, Cooper BG, Hall GL, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. Am J Respir Crit Care Med 2019 Oct 15;200(8):e70-e88 [FREE Full text] [doi: 10.1164/rccm.201908-1590ST] [Medline: 31613151]
14. Bourbeau J, Julien M, Maltais F, Rouleau M, Beaupré A, Bégin R, Chronic Obstructive Pulmonary Disease axis of the Respiratory Network Fonds de la Recherche en Santé du Québec. Reduction of hospital utilization in patients with chronic obstructive pulmonary disease: a disease-specific self-management intervention. Arch Intern Med 2003 Mar 10;163(5):585-591. [doi: 10.1001/archinte.163.5.585] [Medline: 12622605]
15. Muggeridge DJ, Hickson K, Davies AV, Giggins OM, Megson IL, Gorely T, et al. Measurement of Heart Rate Using the Polar OH1 and Fitbit Charge 3 Wearable Devices in Healthy Adults During Light, Moderate, Vigorous, and Sprint-Based Exercise: Validation Study. JMIR Mhealth Uhealth 2021 Mar 25;9(3):e25313 [FREE Full text] [doi: 10.2196/25313] [Medline: 33764310]

Abbreviations

COPD: chronic obstructive pulmonary disease
FEV_1: forced expiratory volume in 1 second
FVC: forced vital capacity
GSM: Global System for Mobile Communications
HRA: Human Research Act
PROM: patient-reported outcome
RCT: randomized controlled trial

Edited by A Kushniruk; submitted 22.06.21; peer-reviewed by R Almeida, E Borycki, A Joseph; comments to author 15.10.21; revised version received 02.11.21; accepted 03.01.22; published 16.02.22.

Please cite as:
Kohlbrenner D, Clarenbach CF, Ivankay A, Zimmerli L, Gross CS, Kuhn M, Brunschwiler T
Multisensory Home-Monitoring in Individuals With Stable Chronic Obstructive Pulmonary Disease and Asthma: Usability Study of the CAir-Desk
JMIR Hum Factors 2022;9(1):e31448
URL: https://humanfactors.jmir.org/2022/1/e31448
doi:10.2196/31448
PMID:35171107

Original Paper

Understanding Patients' Intention to Use Digital Health Apps That Support Postdischarge Symptom Monitoring by Providers Among Patients With Acute Coronary Syndrome: Survey Study

Jinying Chen[1,2], PhD; Jessica G Wijesundara[1], MPH; Gabrielle E Enyim[1], BS; Lisa M Lombardini[1], MS; Ben S Gerber[1], MPH, MD; Thomas K Houston[2], MPH, MD; Rajani S Sadasivam[1], PhD

[1]Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, United States
[2]Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC, United States

Corresponding Author:
Jinying Chen, PhD
Department of Population and Quantitative Health Sciences
University of Massachusetts Chan Medical School
368 Plantation Street
Worcester, MA, 01605
United States
Phone: 1 508 856 6063
Email: jinying.chen@umassmed.edu

Abstract

Background: After hospital discharge, patients with acute coronary syndrome (ACS) often experience symptoms that prompt them to seek acute medical attention. Early evaluation of postdischarge symptoms by health care providers may reduce unnecessary acute care utilization. However, hospital-initiated follow-up encounters are insufficient for timely detection and assessment of symptoms. While digital health tools can help address this issue, little is known about the intention to use such tools in ACS patients.

Objective: This study aimed to assess ACS patients' intention to use digital health apps that support postdischarge symptom monitoring by health care providers and identify patient-perceived facilitators and barriers to app use.

Methods: Using email invitations or phone calls, we recruited ACS patients discharged from a central Massachusetts health care system between December 2020 and April 2021, to participate in the study. Surveys were delivered online or via phone to individual participants. Demographics and access to technology were assessed. The intention to use a symptom monitoring app was assessed using 5-point Likert-type (from strongly agree to strongly disagree) items, such as "If this app were available to me, I would use it." Responses were compared across demographic subgroups and survey delivery methods. Two open-ended questions assessed perceived facilitators and barriers to app use, with responses analyzed using qualitative content analysis.

Results: Among 100 respondents (response rate 8.1%), 45 (45%) completed the survey by phone. The respondents were on average 68 years old (SD 13 years), with 90% (90/100) White, 39% (39/100) women, and 88% (88/100) having access to the internet or a mobile phone. Most participants (65/100, 65%) agreed or strongly agreed that they would use the app, among which 53 (82%) would use the app as often as possible. The percentage of participants with the intention to use the app was 75% among those aged 65-74 years and dropped to 44% among those older than 75 years. The intention to use was higher in online survey respondents (vs phone survey respondents; odds ratio 3.07, 95% CI 1.20-7.88) after adjusting for age and access to technology. The analysis of open-ended questions identified the following 4 main facilitators (motivations): (1) easily reaching providers, (2) accessing or providing information, (3) quickly reaching providers, and (4) consulting providers for symptoms, and the following 4 main barriers: (1) privacy/security concerns, (2) uncomfortable using technology, (3) user-unfriendly app interface, and (4) preference for in-person/phone care.

Conclusions: There was a strong intention to use a symptom monitoring app postdischarge among ACS patients. However, this intent decreased in patients older than 75 years. The survey identified barriers related to technology use, privacy/security, and the care delivery mode. Further research is warranted to determine if such intent translates into app use, and better symptom management and health care quality.

(JMIR Hum Factors 2022;9(1):e34452) doi:10.2196/34452

KEYWORDS

coronary; monitor; elder; health app; symptom; eHealth; mobile health; intention; barrier; facilitator

Introduction

The transition from inpatient care to home is challenging for patients with acute coronary syndrome (ACS) [1-3]. After hospital discharge, ACS patients often experience symptoms that prompt them to seek acute medical attention [2-6]. A large portion of these symptoms are noncardiac [3-7], and could be assessed and managed through close follow-up care in the outpatient setting to reduce unnecessary acute care utilization [3,5-7]. Symptom assessment and management are integral to transitional care [8-13], and are also part of the transitional care management services supported by Medicare [14]. However, hospital-initiated follow-up activities alone may be inadequate to detect symptoms in a timely fashion, as new or worsening symptoms may occur between the initial contact and the follow-up appointment [15]. Intensive transitional care programs offering multiple follow-up phone calls or home visits may better capture patient's symptom episodes [11,12], but providing such thorough contact increases the need for staff resources and time, and can be challenging to scale up.

Digital health tools for symptom monitoring can support timely detection and evaluation of patients' symptoms [16-20], and have been successfully integrated with routine cancer care [16,17,21-23]. Some tools allowed patients to report symptoms frequently or at any time [16,17]. However, in general, evidence about the feasibility and efficacy of using these tools to improve patient outcomes is still limited, especially in patients with ACS. A recent study analyzed data related to using a digital symptom monitoring tool (which allowed patients to self-rate and track their symptoms of fatigue) to enhance a patient-centered care intervention for cardiac rehabilitation [24]. This study found that the enhanced intervention improved patient-reported self-efficacy at 6 months postdischarge, compared with usual care (P=.01). However, only 39% of the patients in the intervention group chose to use the digital health tool.

More research is needed to understand the intention, barriers, and facilitators to digital health symptom monitoring in ACS patients. This is particularly true among older adults (≥65 years old) representative of the ACS population. Older adults have unique barriers in using technology, such as lack of knowledge and confidence, age-related changes or disabilities, and skepticism about the benefits [25,26]. Prior studies showed that most patients, including older adults, are ready to accept digital health tools for monitoring mental health conditions and symptoms, but the intention to use decreased with age [27,28]. Understanding these issues may help improve design, development, and adherence to digital symptom monitoring in ACS patients.

This study aimed to assess ACS patients' intention to use digital health tools that support symptom monitoring by providers after hospital discharge. We conducted a survey, using both close-ended and open-ended questions, to assess the intention to use, the difference in the intention by patient characteristics (eg, age), and the facilitators and barriers of using these tools in this patient population. We also compared the intention to use between 2 survey delivery modes (online vs phone).

Methods

Study Design

We analyzed data collected through a survey using both close-ended and open-ended questions. The survey was delivered using 1 of the 2 modes (online surveys and phone calls) to ensure a balanced sample of participants who are comfortable or are not comfortable with the use of technology (ie, filling online surveys).

Ethics Approval

The study was approved by the Institutional Review Board at the University of Massachusetts Chan Medical School. The ethics approval number (ie, the Institutional Review Board Docket Number) for this study is H00018298. The Institutional Review Board approved the use of informed verbal consent procedures. We obtained verbal informed consent from each participant by email or phone.

Survey

The survey design was informed by prior literature on assessing participants' intention to use digital interventions [29,30]. One researcher (with expertise in health informatics and implementation science) created the initial survey by adapting a subset of validated questions from a survey assessing participants' intention to use mobile apps for COVID-19 symptom monitoring [30]. A cardiologist and 2 research team members (with training in public health and clinical research, respectively) reviewed the survey content and provided feedback on clarifying and simplifying the language of the introduction paragraph, the survey questions, and the response options.

The final survey (Multimedia Appendix 1) included 5 items to assess participants' demographics (age, sex, and race) and access to technology (internet and smartphone), and 5 items (3 close-ended and 2 open-ended questions) related to the intention to use a hypothetical symptom monitoring app. The demographics questions and the open-ended questions were optional. Intention to use the app was assessed using a 5-point Likert-type (from "strongly agree" to "strongly disagree") item (also called the intention-to-use question) as follows: "If this app were available to me, I would use it." Participants who responded "strongly agree," "agree," or "neutral" to this item were prompted to respond to 2 additional items. The first item was a 5-point Likert-type item as follows: "I plan to use this app as often as necessary," with response options ranging from "strongly agree" to "strongly disagree." The second item was multiple-choice as follows: "I'd like the app to be designed as …," with the following 3 options: "mobile app," "web app," and "other." The 2 remaining open-ended questions collected free-text comments on the facilitators (ie, motivations) and barriers to using the app.

Recruitment and Data Collection

We recruited patients from UMass Memorial Health Care, the largest health care system in Central Massachusetts, serving most patients hospitalized with cardiovascular diseases in this region.

Using information from electronic health records (EHRs), we identified adult patients (>18 years old) who were hospitalized for ACS (ICD-10 codes: I24.9, I21, I21.x, I21.xx, and I25.110) between January 2019 and December 2020, as eligible participants. Study data were collected and managed using REDCap electronic data capture tools hosted at the study institution [31,32].

We recruited participants with a 2-stage procedure, using emails and phone calls, respectively. In the first stage (December 2020), we emailed invitations to 782 candidate participants. Once a participant replied to the email to indicate their interest, we sent the online survey via a secure REDCap link to their email address. An unanswered survey was automatically disabled in REDCap 30 days after being sent to the participant. Recruitment stopped after more than 40 participants responded to the online survey.

In the second stage (January 2021 to April 2021), we recruited participants who did not have an email address listed in the EHR via phone calls. Recruitment calls were made to 448 candidate participants until the total number of responses to the survey (from both email and phone recruitment) met the target (N=100). For phone recruitment, we documented the reasons for declining participation. Participants recruited by phone were given the option to complete the survey online (using the same procedure described for stage 1) or via phone. For surveys answered by phone, a research staff member documented participants' verbal responses in REDCap. Each survey participant (for both stages of participant recruitment) was provided a US $10 gift card to compensate for their time.

Research Questions

The following 4 research questions were considered: (1) Do patients have the intention to use the app for symptom monitoring by providers? (Q1); (2) Is there a difference in the intention to use the app for symptom monitoring across subgroups characterized by participants' characteristics, including age and access to technology? (Q2); (3) Is there a difference in the intention to use the app for symptom monitoring between participants responding to the survey online and those responding by phone? (Q3); and (4) What are the main factors that motivate or discourage patients' use of an app for symptom monitoring by providers? (Q4).

Statistical Analyses

Statistical analyses were performed using STATA/IC 15.1 (StataCorp). We first calculated descriptive statistics of participants' characteristics and examined their distributions over the 2 survey delivery modes. We then analyzed the data to answer research questions 1 to 3. We used participants' age information from the EHR, which has greater granularity than the survey responses, for these analyses.

First, we calculated descriptive statistics of participants' responses to the 3 close-ended survey questions related to the intention to use the symptom monitoring app (Q1). Second, we examined the distribution of the intention to use over participants' characteristics and access to technology (Q2). Third, we assessed the associations between survey delivery mode and participants' intention to use the app (Q3), using multivariable logistic regression to adjust for potential confounding factors related to participants' characteristics and access to technology. We identified the confounders based on the literature and the examination of the distribution of participants' characteristics over survey delivery mode ($P<.05$). In addition, we combined access to the internet and access to a smartphone into 1 variable, access to technology, when adjusting for the association analysis because the 2 variables are interdependent (Fisher exact test $P<.001$).

When conducting analyses related to questions 2 and 3, we grouped the 5 response options of the intention-to-use question into 2 categories, with 1 representing "agree" and "strongly agree" and 0 representing the other options. In addition, we assigned numeric values to the 5 response options (1: strongly disagree, 2: disagree, 3: neutral, 4: agree, 5: strongly agree) and presented the summary statistics of the responses.

Qualitative Analyses

To answer research question 4, we analyzed survey responses to the 2 open-ended survey questions through an iterative process using qualitative content analysis. Qualitative content analysis is a research method widely used to analyze written, verbal, or visual communication messages through the systematic coding and identification of themes or patterns [33-35]. Following established techniques [34,35], we carried on the analysis over 3 phases (ie, preparation, organizing, and reporting).

In the preparation phase, GEE (premed student with training in biology, neuroscience, and clinical research) read through the survey responses and assigned initial codes to the responses. JC (with expertise in health informatics and implementation science), JGW (with training in public health and health education), and GEE discussed the initial coding results and created the initial codebook. Using the initial codebook, GEE, JGW, and LML (with training in clinical research and neuroscience) coded all survey responses independently. Codes were assigned to each response (primarily single sentences), and double coding was allowed. The coded responses were discussed among GEE, JGW, LML, and JC to resolve discrepancies, and new codes were added when necessary. This process resulted in the final codebook (Multimedia Appendix 2), with 9 codes (4 categories) for the facilitator question and 8 codes (4 categories) for the barrier question. Based on the coding results, JC segmented survey responses into units that entail a single code. Most segments were single sentences; some were phrases or contained multiple sentences.

In the organizing phase, JC and JGW independently coded the segments using the final codebook. The intercoder agreement was 86% for the facilitator question and 87% for the barrier question. Discrepancies were discussed and resolved between JC and JGW to generate the final coding results.

In the reporting phase, we reported the definitions, frequencies, and representative quotes of codes and summarized key findings [34,35]. We identified the major barriers and facilitators to app use by considering code/category frequency and existing literature on health app use among patients or older adults, and through discussion in the research team. In addition, we compared the most salient facilitators and barriers for the following 2 age groups: younger and older than 65 years of age.

Results

Participant Characteristics

Among 782 patients contacted by email, 59 (7.5%) showed interest in participating in the study, and 48 (81%) of them responded to the survey. Among 448 patients contacted by phone calls, 61 (13.6%) showed interest, and 52 (85%) of them responded to the survey. Overall, the survey response rate was 8.1% (100/1230). There was no difference in age between patients who responded to the survey and patients who did not, including those who did not show interest in participating in the study (67.6 vs 67.7 years, *P*=.94). Of the patients contacted for this study and who did not want to participate, 73 provided reasons for nonparticipation. The common reasons included poor health condition (n=31, 42%), no interest (n=17, 23%), no time (n=11, 15%), and no access or uncomfortable with the use of technology (n=9, 12%).

Among 100 respondents, 45% (ie, 45 of the participants recruited by phone) completed the survey by phone and 55% completed it online. The respondents were on average 68 years old (SD 13 years), with 90% (90/100) White, 39% (39/100) women, and 88% (88/100) reporting having access to the internet or a mobile phone. As shown in Table 1, the rates of access to the internet (*P*<.001) and a smartphone (*P*<.001) were higher in online survey respondents than phone survey respondents. Among the 62 older participants (≥65 years old), 49 (79%) and 41 (66%) reported having access to the internet and a smartphone, respectively.

Table 1. Participant characteristics overall and by the survey delivery mode.

Characteristic	Total (N=100), n (%)	Survey delivery mode, n (%)		*P* value[a]
		Phone (n=45)	Online (n=55)	
Age group				.82
<65 years	38 (38)	16 (36)	22 (40)	
65-74 years	32 (32)	14 (31)	18 (33)	
≥75 years	30 (30)	15 (33)	15 (27)	
Gender				.41
Female	39 (39)	20 (44)	19 (35)	
Male	59 (59)	24 (53)	35 (64)	
Not reported	2 (2)	1 (2)	1 (2)	
Race				>.99
White	90 (90)	39 (87)	51 (93)	
Others	6 (6)	3 (7)	3 (5)	
Not reported	4 (4)	3 (7)	1 (2)	
Has access to the internet				<.001[b]
No	15 (15)	14 (31)	1 (2)	
Yes	85 (85)	31 (69)	54 (98)	
Has a smartphone				<.001[b]
No	25 (25)	19 (42)	6 (11)	
Yes	75 (75)	26 (58)	49 (89)	

[a]Calculated by the Fisher exact test for categorical variables, using complete case analysis (ie, ignoring missing values for gender and race).

[b]Statistically significant (*P*<.05).

Intention to Use the Symptom Monitoring App

All participants (N=100) responded to the intention-to-use survey item, with responses of strongly agree (n=19), agree (n=46), neutral (n=15), disagree (n=15), and strongly disagree (n=5). A total of 74 participants responded to the survey item "I plan to use this app as often as necessary," with responses of strongly agree (n=22), agree (n=35), neutral (n=16), disagree (n=1), and strongly disagree (n=0). Among the 65 (65%) respondents with a positive intention (agree or strongly agree) to use the app, 53 (82%) agreed or strongly agreed that they would use the app as often as possible. Among the 73 respondents to the app design question, 28 (38%) preferred a mobile app, 30 (41%) preferred a web-based app, 14 (19%)

liked both mobile and web-based apps, and 1 (1%) preferred another design (unspecified).

Intention to Use by Patient Characteristics

Among the 62 older participants (≥65 years old), 37 (60%) reported having the intention to use the app. As shown in Table 2, survey respondents aged 75 years or older had a lower rate of intention (ie, agree or strongly agree) to use the app (43%) than those in other age groups (74% for ages under 65 years and 75% for ages 65-74 years; Fisher exact test *P*=.02). There was no difference in the intention to use by gender or race. The rate of the intention to use the app was higher in respondents with access to the internet or a smartphone than those without access (72% vs 17%, *P*<.001).

The mean (Table 2) and median (Multimedia Appendix 3) scores of the intention to use and the distributions of the 5 levels of the intention to use (Multimedia Appendix 3), stratified by participant characteristics, showed similar patterns.

Table 2. Distribution of the intention to use a symptom monitoring app by patient characteristics and the survey delivery mode.

Variable[a]	Response score[b], mean (SD)	Rate of a positive (agree or strongly agree) intention to use the app		
		n/N	%	*P* value[c]
All	3.6 (1.1)	65/100	65	
Age group				.02[d]
<65 years	3.9 (0.8)	28/38	74	
65-74 years	3.7 (1.1)	24/32	75	
≥75 years	3.1 (1.3)	13/30	43	
Gender				>.99
Female	3.6 (1.0)	25/39	64	
Male	3.6 (1.2)	39/59	66	
Race				.66
White	3.6 (1.1)	59/90	66	
Others	3.3 (0.8)	3/6	50	
Has access to technology (internet or a smartphone)				<.001[d]
No	2.2 (1.0)	2/12	17	
Yes	3.8 (1.0)	63/88	72	
Survey delivery mode				.001[d]
Phone	3.1 (1.3)	21/45	47	
Online	4.0 (0.8)	44/55	80	

[a]The gender and race variables had 2 and 4 missing values, respectively.

[b]Scores assigned to the response options were as follows: 1, strongly disagree; 2, disagree; 3, neutral; 4, agree; 5, strongly agree.

[c]Calculated by the Fisher exact test for all the items.

[d]Statistically significant (*P*<.05).

Intention to Use by the Survey Delivery Mode

The rate of a positive intention to use the app (Table 2) was higher in online survey respondents than in phone survey respondents (80% vs 47%, *P*=.001). After adjusting for age and access to technology, the difference remained significant (adjusted odds ratio 3.07, 95% CI 1.20-7.88).

Similarly, the mean (Table 2) and median (Multimedia Appendix 3) scores of the intention to use were higher in online survey respondents (mean 4.0, median 4) than in phone survey respondents (mean 3.1, median 3).

Facilitators and Barriers to Using the App

A total of 84 (84%) participants responded to the facilitator question, for which we identified 73 segments (from 66 participants) that described facilitators. A total of 80 (80%) participants responded to the barrier question, for which we identified 70 segments (from 63 participants) that described barriers. The analyses of these segments identified 9 facilitators or motivations (Figure 1) and 9 barriers (Figure 2). The major facilitators included (1) easily reaching providers, (2) accessing or providing information, (3) quickly reaching providers, and (4) consulting providers for symptoms. We distinguished between barriers 1 and 3, with barrier 1 focusing on convenience in care access (see code definition and more example quotes in Multimedia Appendix 2). The main barriers included (1) privacy/security concerns, (2) uncomfortable using technology, (3) user-unfriendly app interface, and (4) preference for in-person/phone care.

Among participants under 65 years, 87% (33/38) mentioned facilitators to app use, with the most noticeable one being "easily reach providers" (frequency of 14). Among participants aged 65 years or older, 53% (33/62) mentioned facilitators, with the most noticeable one being "access and provide information" (frequency of 8). Among participants under 65, 55% (21/38) mentioned barriers to app use, with the most noticeable one being "lack of timely response" (frequency of 5). Among participants aged 65 years or older, 65% (40/62) mentioned barriers, with the most noticeable one being "uncomfortable with technology" (frequency of 12).

Figure 1. Facilitators to using a symptom monitoring app. Each segment was assigned a single code (ie, facilitator). We have provided an example quote for each code (in parallel to the bars in the figure). More example quotes are provided in Multimedia Appendix 2.

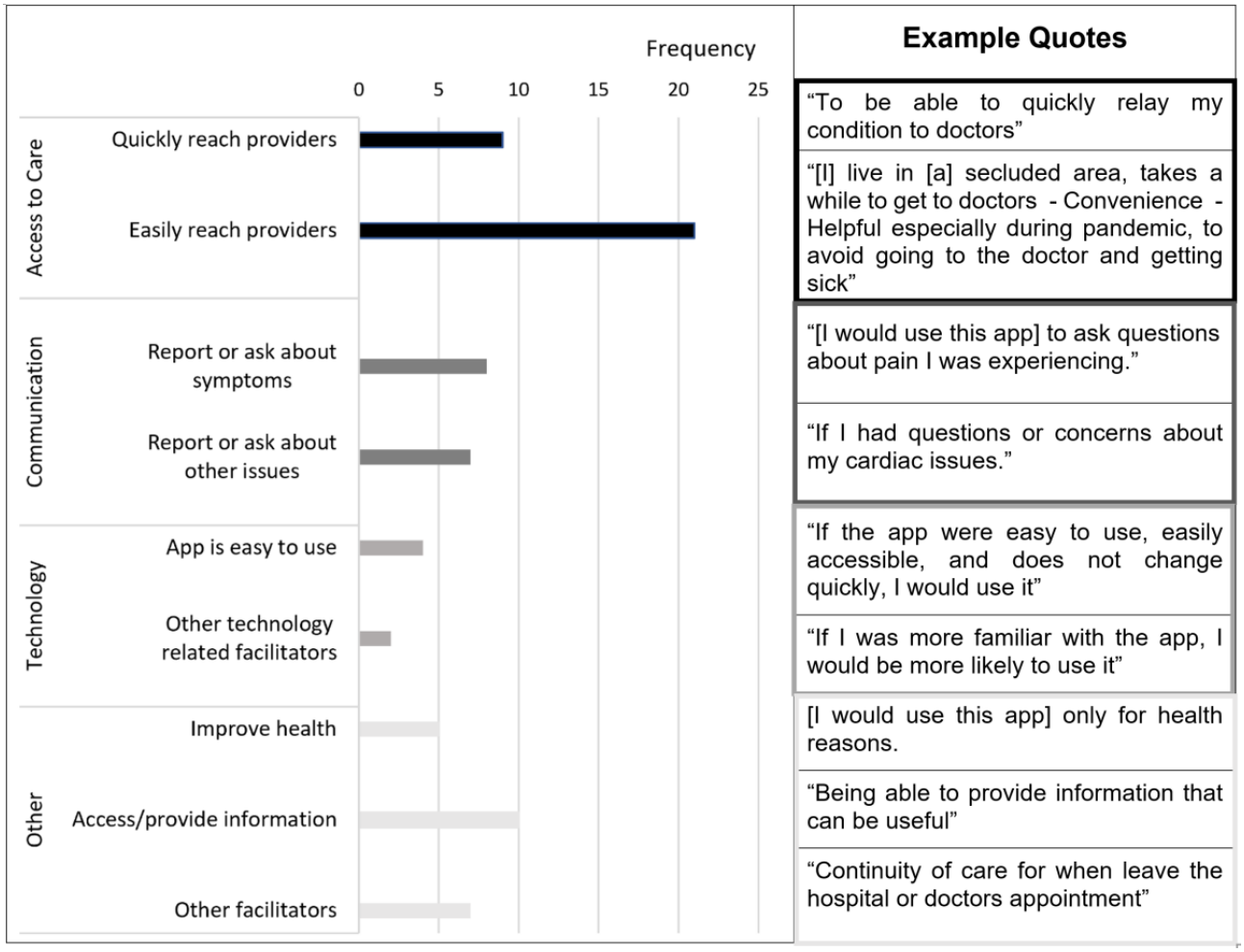
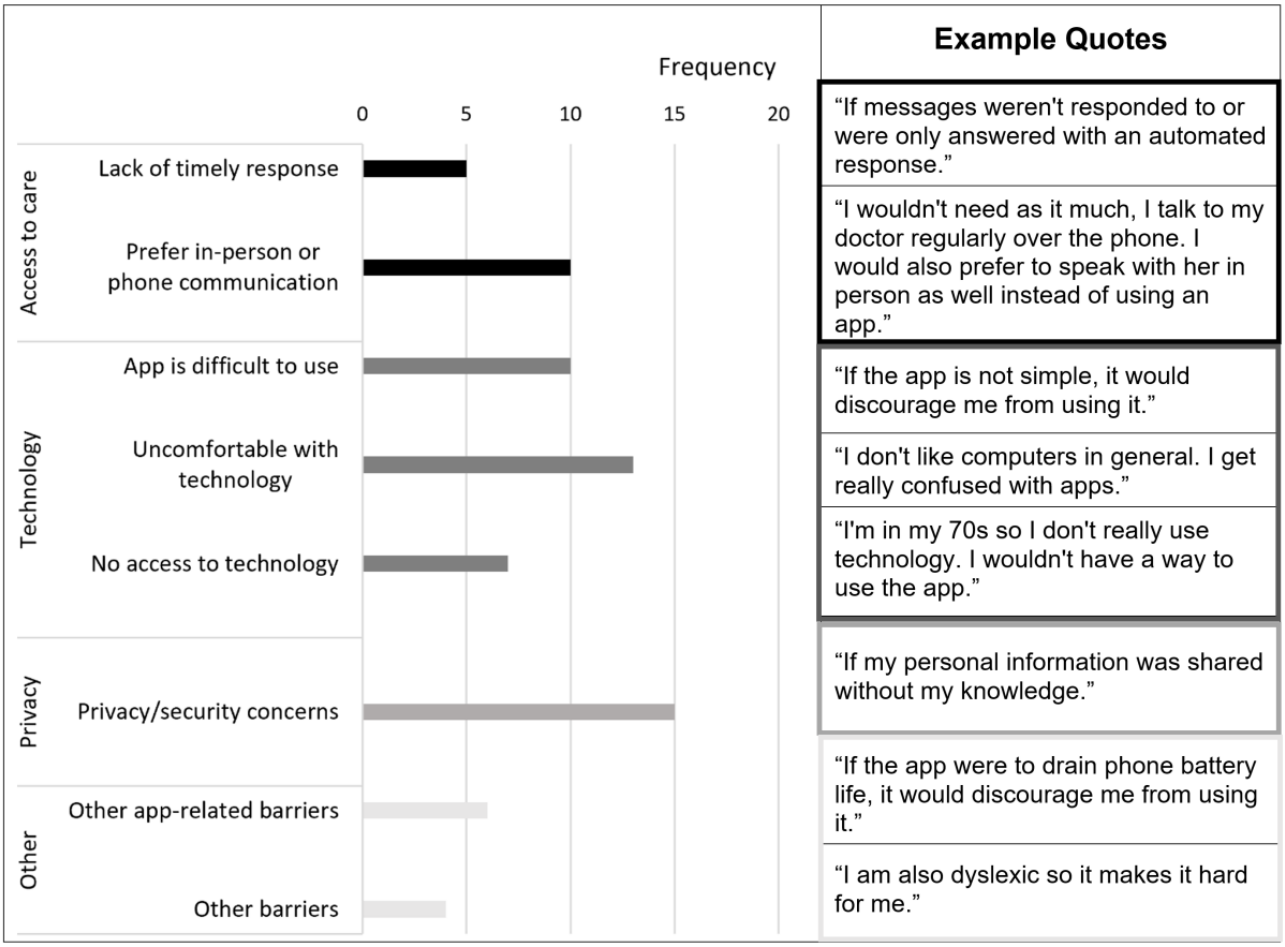


Figure 2. Barriers to using a symptom monitoring app. Each segment was assigned a single code (ie, barrier). We have provided an example quote for each code (in parallel to the bars in the figure). More example quotes are provided in Multimedia Appendix 2.



Discussion

Principal Findings

This is the first study to assess the intention to use a postdischarge symptom monitoring app in ACS patients. We found that most (65/100, 65%) ACS patients had the intention to use an app to monitor and report postdischarge symptoms to providers. Compared with other participants, those aged 75 years or older or lacking access to technology (ie, internet and smartphones) had a lower intention to use the app. Furthermore, phone survey respondents had a lower intention to use the app than online survey respondents. Open-ended survey questions identified important facilitators (Figure 1) and barriers (Figure 2) to using the app in the following 4 domains: access to care, communication, technology, and privacy.

Intention to Use Digital Symptom Monitoring in Older Patients With ACS

Although ACS patients are mostly older adults, we still found a high intention to use the symptom monitoring tool in this population. Specifically, 60% of older participants (≥65 years old) had the intention to use the app. Furthermore, the percentage of participants aged 65-74 years who had the intention to use the app (75%) was as high as that (74%) among younger participants. Our findings are compatible with previous findings on the intention to use health information technology, including symptom monitoring apps, in older adults [27,28,36-41]. For example, prior studies found that 46%-51% of participants older than 60 years would like to use a mobile app to track mental health conditions [27,28]. Other studies also found mobile symptom tracking apps acceptable for older patients with heart failure [38,39], and an app incorporating design features specific to older adults received high usability scores [39]. Similar to prior studies [27,42], we found that older participants had a lower intention to use the app, but we saw this pattern only in participants aged 75 years or older.

Lack of an Email Address in the EHR: A Potential Indicator for a Low Intention to Use Digital Symptom Monitoring

For this study, we intentionally used phone calls to recruit patients who did not have an email address in the EHR. The absence of an email address may imply a lack of email access, infrequent use of email, or less comfort with sending and receiving emails. Most of these participants (ie, those without an email address in the EHR) chose to complete the survey over the phone and had a lower intention to use a symptom monitoring app, even after adjusting for age and access to technology. This suggests that a lack of an email address itself may be a useful predictor and provide meaningful information for health care teams making decisions about remote symptom monitoring postdischarge. In the future, this information (ie, lack of an email address in the EHR) can be used to purposefully

sample key informants to help design and user test symptom monitoring apps and identify patients who may need greater training and support in app use.

Patient-Perceived Facilitators and Barriers to Using Digital Symptom Monitoring

This study also identified important facilitators and barriers to using a symptom monitoring app in ACS patients. Prior studies found that perceived usefulness significantly influenced the intention to use medical apps in older patients [40,41,43,44]. Similarly, we found that the facilitators or motivations to using a symptom monitoring app mainly were related to perceived usefulness of the app, such as reaching health providers easily, accessing and providing health information, and consulting with providers regarding symptom management. The major barrier identified was patients' concerns with privacy and security. This is common with digital health interventions and needs to be addressed from the perspectives of both the app and the users [45-48]. In addition to following the regulations and incorporating standard security features in app design [47,48], it is important to assess user opinions on desired privacy and security features in their local context [46,49]. In this study, we found that ACS patients were concerned about who will access their health information and the disclosure of their health information to a third party without their knowledge and authorization. Using hospital-authorized apps, clearly communicating with patients an app's privacy statement, and providing options for choosing which information to disclose with whom may reduce this barrier. Similar to prior studies [25,26], we found that the most notable barrier for using the symptom app in older (≥65 years old) ACS patients is being uncomfortable using technology. Patient-centered app design, in-hospital training for app use, and app use support from caregivers may help reduce the barriers [50].

Previous studies found that patients sometimes have challenges in deciding when to use an app to report symptoms. For example, patients sometimes reported urgent issues via secure messaging services designed for communicating nonurgent issues [51-53]. In addition, prior studies found that ACS patients were more stressful about certain symptoms and 15% of patients developed stress disorder symptoms after ACS [54,55]. It is likely that some patients would unnecessarily seek acute care when experiencing nonurgent symptoms [56]. In this study, we did not find these issues to be a theme when analyzing patient-reported barriers to app use. However, it is important to communicate with patients about the appropriate use of a symptom monitoring app and how frequently providers would review or respond to patient reporting. Patient education on how to assess the severity of symptoms, for example, identifying typical ACS symptoms that need urgent care, is also relevant and may improve health care utilization.

Implications on App Design and Development

Whether an intention to use a digital health app can translate into real use depends on many factors, such as app design and implementation strategies to support app use. In addition to general app design principles (eg, secure and easy to use), this study suggests additional considerations in app development for ACS patients. Specifically, we found that older age and lack of access to technology were associated with a low intention to use the app, and the most common barrier to app use in older adults was being uncomfortable using technology. This suggests that a multimodal strategy may be more effective in engaging these patients. For those who have nonsmart phones or are less comfortable using apps, text messaging may serve as an additional communication channel. Alternatively, app design may allow for the involvement of family members or caregivers in symptom tracking. In addition, accessible design principles for older adults may be incorporated by including a consistent and simple interface, making the most essential functionalities readily visible and available, and making it easier to "undo" an unintended action [57,58]. A co-creation approach that engages older patients in all stages of app development and user testing is also important for improving app adoption and user experience [59,60].

In this study, we also found that patients were motivated to use an app to easily reach providers. Therefore, the app should allow providers to easily access symptom reports, triage symptoms, and respond to patient symptoms and concerns. It is also critical to engage providers in all phases of app design and testing. App adoption will need to address how to integrate information from the app into the EHR, and assess the impact of the app on provider burden and clinical workflow [61,62].

Limitations

Our sample was relatively small and from a health care system in 1 state, and most participants were non-Hispanic White. Therefore, our findings may not be generalizable to other settings. Constrained by the format of a survey study, participants' responses to the open-ended survey questions were typically short and lacked detailed information about the contextual factors related to the perceived facilitators and barriers. We interpret these qualitative results based on the existing literature. In-depth qualitative studies are warranted to better understand certain barriers, such as the preference for in-person care and phone communication.

Conclusions

We found a strong intention of using a symptom monitoring app postdischarge among ACS patients. However, this intent was lower in patients aged 75 years or older. Our survey identified barriers related to privacy and security, technology use, and the care delivery mode. Using hospital-authorized apps and in-hospital training may reduce the barriers. Further research is warranted to determine if such intent translates into app use, and better symptom management and health care quality.

Acknowledgments

This work was supported by the National Heart, Lung and Blood Institute (grant 1K12HL138049-3). TKH, RSS, and JC also received funding support from the National Cancer Institute (grant 1P50CA244693). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

JC and RSS conceived this study. JC designed the study with inputs from RSS, JGW, and LML. LML and GEE recruited and surveyed the participants. JC analyzed the quantitative survey data. GEE, JGW, LML, and JC coded responses to open-ended survey questions, and JC and JGW finalized the qualitative analyses. RSS, BSG, and TKH provided expertise in digital health research and critical intellectual inputs to data analysis and result interpretation. JC drafted the manuscript. JGW, GEE, LML, and BSG contributed to paper writing. RSS, BSG, and TKH provided important feedback for paper revision. All authors reviewed and provided feedback for the manuscript and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey to assess the intention to use a symptom monitoring app.
[DOCX File , 18 KB - humanfactors_v9i1e34452_app1.docx]

Multimedia Appendix 2
Codebook and example quotes for facilitators and barriers to using a symptom monitoring app.
[DOCX File , 28 KB - humanfactors_v9i1e34452_app2.docx]

Multimedia Appendix 3
Intention to use a symptom monitoring app, stratified by patient characteristics and the survey delivery mode.
[DOCX File , 24 KB - humanfactors_v9i1e34452_app3.docx]

References

1. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. Circulation 2020 Mar 03;141(9):e139-e596 [FREE Full text] [doi: 10.1161/CIR.0000000000000757] [Medline: 31992061]
2. Dharmarajan K, Hsieh AF, Lin Z, Bueno H, Ross JS, Horwitz LI, et al. Diagnoses and timing of 30-day readmissions after hospitalization for heart failure, acute myocardial infarction, or pneumonia. JAMA 2013 Jan 23;309(4):355-363 [FREE Full text] [doi: 10.1001/jama.2012.216476] [Medline: 23340637]
3. Wasfy JH, Strom JB, O'Brien C, Zai AH, Luttrell J, Kennedy KF, et al. Causes of short-term readmission after percutaneous coronary intervention. Circ Cardiovasc Interv 2014 Feb;7(1):97-103. [doi: 10.1161/CIRCINTERVENTIONS.113.000988] [Medline: 24425587]
4. Southern DA, Ngo J, Martin B, Galbraith PD, Knudtson ML, Ghali WA, et al. Characterizing types of readmission after acute coronary syndrome hospitalization: implications for quality reporting. J Am Heart Assoc 2014 Sep 18;3(5):e001046 [FREE Full text] [doi: 10.1161/JAHA.114.001046] [Medline: 25237046]
5. Kwok CS, Shah B, Al-Suwaidi J, Fischman DL, Holmvang L, Alraies C, et al. Timing and Causes of Unplanned Readmissions After Percutaneous Coronary Intervention: Insights From the Nationwide Readmission Database. JACC Cardiovasc Interv 2019 Apr 22;12(8):734-748 [FREE Full text] [doi: 10.1016/j.jcin.2019.02.007] [Medline: 30928446]
6. Shah M, Patil S, Patel B, Agarwal M, Davila CD, Garg L, et al. Causes and Predictors of 30-Day Readmission in Patients With Acute Myocardial Infarction and Cardiogenic Shock. Circ Heart Fail 2018 Apr;11(4):e004310. [doi: 10.1161/CIRCHEARTFAILURE.117.004310] [Medline: 29618581]
7. Iribarne A, Chang H, Alexander JH, Gillinov AM, Moquete E, Puskas JD, et al. Readmissions after cardiac surgery: experience of the National Institutes of Health/Canadian Institutes of Health research cardiothoracic surgical trials network. Ann Thorac Surg 2014 Oct;98(4):1274-1280 [FREE Full text] [doi: 10.1016/j.athoracsur.2014.06.059] [Medline: 25173721]
8. Enderlin CA, McLeskey N, Rooker JL, Steinhauser C, D'Avolio D, Gusewelle R, et al. Review of current conceptual models and frameworks to guide transitions of care in older adults. Geriatr Nurs 2013;34(1):47-52. [doi: 10.1016/j.gerinurse.2012.08.003] [Medline: 23122908]
9. Medicare Program; CY 2020 Revisions to Payment Policies Under the Physician Fee Schedule and Other Changes to Part B Payment Policies; Medicare Shared Savings Program Requirements; Medicaid Promoting Interoperability Program Requirements for Eligible Professionals; Establishment of an Ambulance Data Collection System; Updates to the Quality

Payment Program; Medicare Enrollment of Opioid Treatment Programs and Enhancements to Provider Enrollment Regulations Concerning Improper Prescribing and Patient Harm; and Amendments to Physician Self-Referral Law Advisory Opinion Regulations Final Rule; and Coding and Payment for Evaluation and Management, Observation and Provision of Self-Administered Esketamine Interim Final Rule. Centers for Medicare & Medicaid Services. 2019. URL: https://tinyurl.com/2p8zk663 [accessed 2022-02-14]
10. Jack B, Paasche-Orlow M, Mitchell S, Forsythe S, Martin J, Brach C. An overview of the Re-Engineered Discharge (RED) Toolkit (Prepared by Boston University under Contract No. HHSA290200600012i). Boston University. Rockville, MD: Agency for Healthcare Research and Quality; 2012. URL: https://www.bu.edu/fammed/projectred/newtoolkit/ProjectRED-tool1-overview.pdf [accessed 2022-02-14]
11. Coleman EA, Parry C, Chalmers S, Min S. The care transitions intervention: results of a randomized controlled trial. Arch Intern Med 2006 Sep 25;166(17):1822-1828. [doi: 10.1001/archinte.166.17.1822] [Medline: 17000937]
12. Hirschman KB, Shaid E, McCauley K, Pauly MV, Naylor MD. Continuity of Care: The Transitional Care Model. Online J Issues Nurs 2015 Sep 30;20(3):1 [FREE Full text] [Medline: 26882510]
13. Burke RE, Guo R, Prochazka AV, Misky GJ. Identifying keys to success in reducing readmissions using the ideal transitions in care framework. BMC Health Serv Res 2014 Sep 23;14:423 [FREE Full text] [doi: 10.1186/1472-6963-14-423] [Medline: 25244946]
14. Bloink J, Adler KG. Transitional care management services: new codes, new requirements. Fam Pract Manag 2013;20(3):12-17 [FREE Full text] [Medline: 23939734]
15. Felix HC, Seaberg B, Bursac Z, Thostenson J, Stewart MK. Why do patients keep coming back? Results of a readmitted patient survey. Soc Work Health Care 2015;54(1):1-15 [FREE Full text] [doi: 10.1080/00981389.2014.966881] [Medline: 25588093]
16. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. CA Cancer J Clin 2012;62(5):337-347 [FREE Full text] [doi: 10.3322/caac.21150] [Medline: 22811342]
17. Jensen RE, Snyder CF, Abernethy AP, Basch E, Potosky AL, Roberts AC, et al. Review of electronic patient-reported outcomes systems used in cancer clinical care. J Oncol Pract 2014 Jul;10(4):e215-e222 [FREE Full text] [doi: 10.1200/JOP.2013.001067] [Medline: 24301843]
18. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and Analysis of Existing Mobile Phone Apps to Support Heart Failure Symptom Monitoring and Self-Care Management Using the Mobile Application Rating Scale (MARS). JMIR Mhealth Uhealth 2016 Jun 14;4(2):e74 [FREE Full text] [doi: 10.2196/mhealth.5882] [Medline: 27302310]
19. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. J Psychiatr Res 2018 Dec;107:73-78. [doi: 10.1016/j.jpsychires.2018.10.006] [Medline: 30347316]
20. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul 11;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
21. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. J Clin Oncol 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: 10.1200/JCO.2015.63.0830] [Medline: 26644527]
22. Denis F, Basch E, Septans A, Bennouna J, Urban T, Dueck AC, et al. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. JAMA 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: 10.1001/jama.2018.18085] [Medline: 30667494]
23. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
24. Wolf A, Fors A, Ulin K, Thorn J, Swedberg K, Ekman I. An eHealth Diary and Symptom-Tracking Tool Combined With Person-Centered Care for Improving Self-Efficacy After a Diagnosis of Acute Coronary Syndrome: A Substudy of a Randomized Controlled Trial. J Med Internet Res 2016 Feb 23;18(2):e40 [FREE Full text] [doi: 10.2196/jmir.4890] [Medline: 26907584]
25. Gitlow L. Technology Use by Older Adults and Barriers to Using Technology. Physical & Occupational Therapy In Geriatrics 2014 Aug 12;32(3):271-280. [doi: 10.3109/02703181.2014.946640]
26. Vaportzis E, Clausen MG, Gow AJ. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. Front Psychol 2017 Oct 04;8:1687 [FREE Full text] [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
27. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014 Jan 21;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]

28. Torous J, Chan SR, Yee-Marie Tan S, Behrens J, Mathew I, Conrad EJ, et al. Patient Smartphone Ownership and Interest in Mobile Apps to Monitor Symptoms of Mental Health Conditions: A Survey in Four Geographically Distinct Psychiatric Clinics. JMIR Ment Health 2014;1(1):e5 [FREE Full text] [doi: 10.2196/mental.4004] [Medline: 26543905]
29. van Velsen L, van der Geest T, van de Wijngaert L, van den Berg S, Steehouder M. Personalization has a Price, Controllability is the Currency: Predictors for the Intention to use Personalized eGovernment Websites. Journal of Organizational Computing and Electronic Commerce 2015 Feb 05;25(1):76-97. [doi: 10.1080/10919392.2015.990782] [Medline: 26281194]
30. Jansen-Kosterink S, Hurmuz M, den Ouden M, van Velsen L. Predictors to use mobile apps for monitoring COVID-19 symptoms and contact tracing: A survey among Dutch citizens. medRxiv. 2020. URL: https://www.medrxiv.org/content/10.1101/2020.06.02.20113423v1 [accessed 2022-02-14]
31. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
32. 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]
33. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
34. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
35. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
36. Parker SJ, Jessel S, Richardson JE, Reid MC. Older adults are mobile too!Identifying the barriers and facilitators to older adults' use of mHealth for pain management. BMC Geriatr 2013 May 06;13:43 [FREE Full text] [doi: 10.1186/1471-2318-13-43] [Medline: 23647949]
37. Hung L, Lyons JG, Wu C. Health information technology use among older adults in the United States, 2009-2018. Curr Med Res Opin 2020 May;36(5):789-797. [doi: 10.1080/03007995.2020.1734782] [Medline: 32096650]
38. Portz JD, Vehovec A, Dolansky MA, Levin JB, Bull S, Boxer R. The Development and Acceptability of a Mobile Application for Tracking Symptoms of Heart Failure Among Older Adults. Telemed J E Health 2018 Feb 12;24(2):161-165 [FREE Full text] [doi: 10.1089/tmj.2017.0036] [Medline: 28696832]
39. Reading Turchioe M, Grossman LV, Baik D, Lee CS, Maurer MS, Goyal P, et al. Older Adults Can Successfully Monitor Symptoms Using an Inclusively Designed Mobile Application. J Am Geriatr Soc 2020 Jun 10;68(6):1313-1318 [FREE Full text] [doi: 10.1111/jgs.16403] [Medline: 32157679]
40. Askari M, Klaver NS, van Gestel TJ, van de Klundert J. Intention to use Medical Apps Among Older Adults in the Netherlands: Cross-Sectional Study. J Med Internet Res 2020 Sep 04;22(9):e18080 [FREE Full text] [doi: 10.2196/18080] [Medline: 32624465]
41. de Veer AJE, Peeters JM, Brabers AEM, Schellevis FG, Rademakers JJDJM, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15:103 [FREE Full text] [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
42. Sohn A, Speier W, Lan E, Aoki K, Fonarow G, Ong M, et al. Assessment of Heart Failure Patients' Interest in Mobile Health Apps for Self-Care: Survey Study. JMIR Cardio 2019 Oct 29;3(2):e14332 [FREE Full text] [doi: 10.2196/14332] [Medline: 31758788]
43. Cajita MI, Hodgson NA, Budhathoki C, Han H. Intention to Use mHealth in Older Adults With Heart Failure. J Cardiovasc Nurs 2017 Feb 28;32(6):E1-E7 [FREE Full text] [doi: 10.1097/JCN.0000000000000401] [Medline: 28248747]
44. Zhang Y, Liu C, Luo S, Xie Y, Liu F, Li X, et al. Factors Influencing Patients' Intentions to Use Diabetes Management Apps Based on an Extended Unified Theory of Acceptance and Use of Technology Model: Web-Based Survey. J Med Internet Res 2019 Aug 13;21(8):e15023 [FREE Full text] [doi: 10.2196/15023] [Medline: 31411146]
45. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
46. Zhou L, Bao J, Watzlaf V, Parmanto B. Barriers to and Facilitators of the Use of Mobile Health Apps From a Security Perspective: Mixed-Methods Study. JMIR Mhealth Uhealth 2019 Apr 16;7(4):e11223 [FREE Full text] [doi: 10.2196/11223] [Medline: 30990458]
47. Arora S, Yttri J, Nilse W. Privacy and Security in Mobile Health (mHealth) Research. Alcohol Res 2014;36(1):143-151 [FREE Full text] [Medline: 26259009]
48. Morera EP, de la Torre Díez I, Garcia-Zapirain B, López-Coronado M, Arambarri J. Security Recommendations for mHealth Apps: Elaboration of a Developer's Guide. J Med Syst 2016 Jun;40(6):152. [doi: 10.1007/s10916-016-0513-6] [Medline: 27147515]

49. Atienza AA, Zarcadoolas C, Vaughon W, Hughes P, Patel V, Chou WS, et al. Consumer Attitudes and Perceptions on mHealth Privacy and Security: Findings From a Mixed-Methods Study. J Health Commun 2015;20(6):673-679. [doi: 10.1080/10810730.2015.1018560] [Medline: 25868685]
50. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. J Am Med Inform Assoc 2018 Aug 01;25(8):1080-1088 [FREE Full text] [doi: 10.1093/jamia/ocy052] [Medline: 29788380]
51. North F, Crane SJ, Stroebel RJ, Cha SS, Edell ES, Tulledge-Scheitel SM. Patient-generated secure messages and eVisits on a patient portal: are patients at risk? J Am Med Inform Assoc 2013;20(6):1143-1149 [FREE Full text] [doi: 10.1136/amiajnl-2012-001208] [Medline: 23703826]
52. Lanham HJ, Leykum LK, Pugh JA. Examining the Complexity of Patient-Outpatient Care Team Secure Message Communication: Qualitative Analysis. J Med Internet Res 2018 Jul 11;20(7):e218 [FREE Full text] [doi: 10.2196/jmir.9269] [Medline: 29997107]
53. Shimada SL, Petrakis BA, Rothendler JA, Zirkle M, Zhao S, Feng H, et al. An analysis of patient-provider secure messaging at two Veterans Health Administration medical centers: message content and resolution through secure messaging. J Am Med Inform Assoc 2017 Sep 01;24(5):942-949 [FREE Full text] [doi: 10.1093/jamia/ocx021] [Medline: 28371896]
54. Ayers S, Copland C, Dunmore E. A preliminary study of negative appraisals and dysfunctional coping associated with post-traumatic stress disorder symptoms following myocardial infarction. Br J Health Psychol 2009 Sep;14(Pt 3):459-471. [doi: 10.1348/135910708X349343] [Medline: 18789186]
55. Edmondson D, Rieckmann N, Shaffer JA, Schwartz JE, Burg MM, Davidson KW, et al. Posttraumatic stress due to an acute coronary syndrome increases risk of 42-month major adverse cardiac events and all-cause mortality. J Psychiatr Res 2011 Dec;45(12):1621-1626 [FREE Full text] [doi: 10.1016/j.jpsychires.2011.07.004] [Medline: 21807378]
56. Kwok CS, Wong CW, Shufflebotham H, Brindley L, Fatima T, Shufflebotham A, et al. Early Readmissions After Acute Myocardial Infarction. Am J Cardiol 2017 Sep 01;120(5):723-728. [doi: 10.1016/j.amjcard.2017.05.049] [Medline: 28728745]
57. Fisk A, Czaja S, Rogers W, Charness N, Sharit J. Designing for Older Adults: Principles and Creative Human Factors Approaches. Boca Raton: CRC Press; 2019.
58. Farage MA, Miller KW, Ajayi F, Hutchins D. Design principles to accommodate older adults. Glob J Health Sci 2012 Feb 29;4(2):2-25 [FREE Full text] [doi: 10.5539/gjhs.v4n2p2] [Medline: 22980147]
59. Mansson L, Wiklund M, Öhberg F, Danielsson K, Sandlund M. Co-Creation with Older Adults to Improve User-Experience of a Smartphone Self-Test Application to Assess Balance Function. Int J Environ Res Public Health 2020 May 26;17(11):3768 [FREE Full text] [doi: 10.3390/ijerph17113768] [Medline: 32466484]
60. Mannheim I, Schwartz E, Xi W, Buttigieg SC, McDonnell-Naughton M, Wouters EJM, et al. Inclusion of Older Adults in the Research and Design of Digital Technology. Int J Environ Res Public Health 2019 Oct 02;16(19):3718 [FREE Full text] [doi: 10.3390/ijerph16193718] [Medline: 31581632]
61. Mishuris RG, Yoder J, Wilson D, Mann D. Integrating data from an online diabetes prevention program into an electronic health record and clinical workflow, a design phase usability study. BMC Med Inform Decis Mak 2016 Jul 11;16:88 [FREE Full text] [doi: 10.1186/s12911-016-0328-x] [Medline: 27401606]
62. Borycki E, Kushniruk A, Nohr C, Takeda H, Kuwata S, Carvalho C, et al. Usability Methods for Ensuring Health Information Technology Safety: Evidence-Based Approaches. Contribution of the IMIA Working Group Health Informatics for Patient Safety. Yearb Med Inform 2013;8:20-27. [Medline: 23974544]

Abbreviations

ACS: acute coronary syndrome
EHR: electronic health record

Edited by A Kushniruk; submitted 24.10.21; peer-reviewed by M Stemmer, H Gandhi; comments to author 05.12.21; revised version received 17.12.21; accepted 19.12.21; published 07.03.22.

Please cite as:
Chen J, Wijesundara JG, Enyim GE, Lombardini LM, Gerber BS, Houston TK, Sadasivam RS
Understanding Patients' Intention to Use Digital Health Apps That Support Postdischarge Symptom Monitoring by Providers Among Patients With Acute Coronary Syndrome: Survey Study
JMIR Hum Factors 2022;9(1):e34452
URL: https://humanfactors.jmir.org/2022/1/e34452
doi:10.2196/34452
PMID:35254269

Original Paper

The Acceptability of Virtual Characters as Social Skills Trainers: Usability Study

Hiroki Tanaka[1], PhD; Satoshi Nakamura[2], PhD

[1]Division of Information Science, Nara Institute of Science and Technology, Ikoma-shi, Japan

[2]Data Science Center, Nara Institute of Science and Technology, Ikoma-shi, Japan

Corresponding Author:
Hiroki Tanaka, PhD
Division of Information Science
Nara Institute of Science and Technology
Takayamacho 8916-5
Ikoma-shi, 630-0192
Japan
Phone: 81 9076493408
Email: hiroki-tan@is.naist.jp

Abstract

Background: Social skills training by human trainers is a well-established method to provide appropriate social interaction skills and strengthen social self-efficacy. In our previous work, we attempted to automate social skills training by developing a virtual agent that taught social skills through interaction. Previous research has not investigated the visual design of virtual agents for social skills training. Thus, we investigated the effect of virtual agent visual design on automated social skills training.

Objective: The 3 main purposes of this research were to investigate the effect of virtual agent appearance on automated social skills training, the relationship between acceptability and other measures (eg, likeability, realism, and familiarity), and the relationship between likeability and individual user characteristics (eg, gender, age, and autistic traits).

Methods: We prepared images and videos of a virtual agent, and 1218 crowdsourced workers rated the virtual agents through a questionnaire. In designing personalized virtual agents, we investigated the acceptability, likeability, and other impressions of the virtual agents and their relationship to individual characteristics.

Results: We found that there were differences between the virtual agents in all measures (P<.001). A female anime-type virtual agent was rated as the most likeable. We also confirmed that participants' gender, age, and autistic traits were related to their ratings.

Conclusions: We confirmed the effect of virtual agent design on automated social skills training. Our findings are important in designing the appearance of an agent for use in personalized automated social skills training.

(JMIR Hum Factors 2022;9(1):e35358) doi:10.2196/35358

KEYWORDS

social skills training; virtual agent design; virtual assistant; virtual trainer; chatbot; acceptability; realism; virtual agent; simulation; social skill; social interaction; design; training; crowdsourcing

Introduction

Social skills training is a method widely applied to help people who lack social skills. It is used in medical hospitals, employment support facilities, workplaces, schools, and various other institutions [1]. Social skills training is generally conducted by a human trainer to promote appropriate social interaction skills and strengthen social self-efficacy [2]. The Bellack method (or step-by-step social skills training) is a well-structured and widely used evidence-based approach [1]. It is a cognitive behavioral approach to social skills training inspired by the 5 core principles of social learning theory: modeling, shaping, reinforcement, overlearning, and generalization [3]. The Bellack method defines the social skills training framework and its 4 basic skills: expressing positive feelings, listening to others, making requests, and expressing unpleasant feelings. These skills are beneficial for all people (not only those with autistic traits or schizophrenia) [1]. In particular, autism spectrum disorder (ASD) is a spectrum condition [4], meaning it has a broad range of characteristics, from mild to severe. Using

computer agents in social skills training is motivated by the fact that even though some people with high-functioning autism experience difficulty during social communication, they also show good or even superior systemizing skills [5]. Systemizing is the drive to analyze or build systems and understand and predict behavior in terms of underlying rules and regularities. The use of systematic computer-based training for people who need to improve their social skills has the following benefits: (1) it uses a computerized environment that is predictable, consistent, and free from social demands; (2) users can work at their own pace and level of understanding; (3) training can be repeated until the goal is achieved; and (4) interest and motivation can be maintained through computerized rewards. It may also be easier for those who suffer from social difficulties to use computer agents than to directly interact with humans [6]. A past paper suggested that people with social difficulties such as ASD feel safer and more comfortable in virtual interactions than in interactions with actual people [7].

We and other research groups have been conducting studies to automate social skills training using virtual agents, and this work has led to the development of automatic social skills training [8-12] that by design resembles human-led social skills training [10]. The use of conversational agents in health care was reviewed by Tudor Car et al and Milne-Ives et al [13,14]. Among types of conversational agents, our system includes video modeling of human behavior, real-time behavior recognition, and feedback. We previously confirmed the effectiveness of this training in children and adults with ASD and in the general population. The automated social skills training agent plays 2 roles: as a trainer and as a listener. We confirmed that the system was more effective in training social skills than the traditional methods of reading books or watching videos of role models, and that talking to a 3D virtual agent made users feel more comfortable and less tense than talking to a human [15]. Automated social skills training targets various populations, from children to adult men and women, as well as those with ASD or schizophrenia [1]. However, visual designs of virtual agents, and what kind of design is more favored or more accepted, has not yet been investigated. A previous study showed that the quality of the therapeutic alliance (ie, the level of rapport and trust) is a reliable predictor of positive clinical outcomes independent of the approach to psychotherapy (including social skills training [1] and cognitive behavioral therapy [16]) or the specific outcome measure [17]. For automatic social skills training to be adopted and accepted by individuals, detailed investigation is necessary. In this study, we focus on comparing virtual agents with varying visual designs, rather than comparing humans and robots [18] for assistive technology [19], because we consider that the design of virtual agents is easier to create and modify.

The visual design of the virtual agent in social skills training has been previously investigated, although not exhaustively. For example, Hoque et al [12] paired male participants with a male coach and female participants with a female coach in order to minimize gender-based variability in behavior. By contrast, Tanaka et al [10,15] did not consider the agent's gender (they used only a female design). Previous studies have used various virtual agent designs for different tasks and compared their appearance and behavior [20-22], realism [23], intensity in dialogue scenarios, and the appropriateness of body and eye proportions [24]. Past studies have also created a voice designed for the elderly [25] and have examined the impact of gender and race on users' self-efficacy [26]. Troncone et al [27] discussed seniors' psychological perspectives in terms of the model of acceptance and associated factors. Our study applies these findings and rating measures to investigate the design of our virtual agents, aiming to create a more favorable and acceptable design for automated social skills training. To the best of our knowledge, previous work has not investigated the visual design of virtual agents for automated social skills training, the relationship between acceptability and other measures, and the relationship between likeability and individual user characteristics.

This study set Japanese adults as our target users. We prepared a variety of new virtual agent designs for social skills training and evaluated them with multiple items on a questionnaire: their acceptability as a trainer; their acceptability as a listener; their realism, familiarity, trustworthiness, and eeriness; the likeability of their face, eyes, hair, perceived age, and voice; and their overall impression. These criteria were chosen with reference to the studies of Esposito et al [20,21,25] and Ring et al [24]. We followed their statistical analysis framework and investigated the appearance of 3D characters in the context of automated social skills training. First, we evaluated virtual agent visual design, particularly realism. Previous work has showed that serious tasks, such as medical diagnosis, require realistic agents; on the other hand, anime-like agents are more suited to social chitchat-like dialogue systems [24]. We hypothesized that anime-like characters would be preferable and more accepted for automated social skills training since such training requires friendly characteristics to maintain participant safety, and because agents play 2 roles: as trainers and as listeners. In addition, realism is affected by the "uncanny valley" phenomenon, with the most unrealistic character often being rated as the most acceptable [23]. We hypothesized that we would find that the uncanny valley also applies to automated social skills training agents. Second, to examine new factors that correlate to acceptability, we quantified the relationships between acceptability and other measures. We hypothesized that these questionnaire items would be highly correlated with each other [23]. Finally, we investigated the differences in preference for virtual agent design by considering individual users' gender, age, and autistic traits in order to enable personalized automated social skills training. The three main research problems were (1) to investigate the visual appearance of virtual agents for automated social skills training; (2) to investigate the relationship between acceptability and other measures (eg, likeability, face, voice, realism, and familiarity); and (3) to investigate the relationship between acceptability and the individual characteristics of the user (ie, gender, age, and autistic traits).

This paper is an extension of conference proceedings [28] in which we reported on the visual design of characters. This paper adds an analysis of realism and includes a greater number of participants. We created new agents and videos and evaluated their realism. We also investigated whether people with high

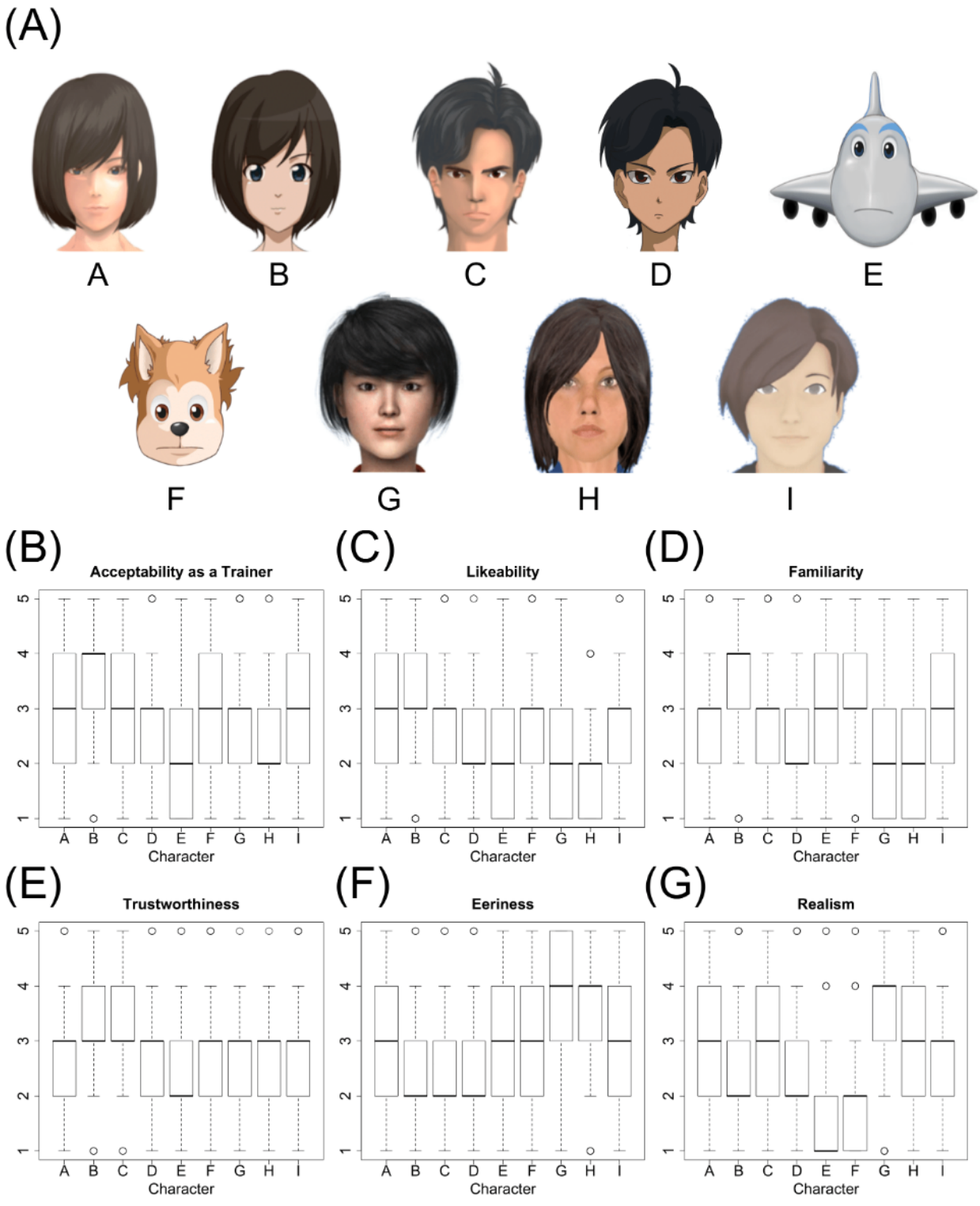
or low autistic traits rated the likeability of virtual agents differently depending on the realism of the agent. We also analyzed the correlation matrix between all questionnaire items in order to confirm correlations between acceptability and other measures. Finally, this paper discusses and summarizes findings from a series of experiments.

Methods

Visual Design of Virtual Agents

We first prepared an illustration of a virtual agent, as shown in Figure 1. The virtual agents were designed by a company specializing in Japanese animation.

Figure 1. Images of the 9 virtual characters and representative measures collected from data set 1 and data set 2.



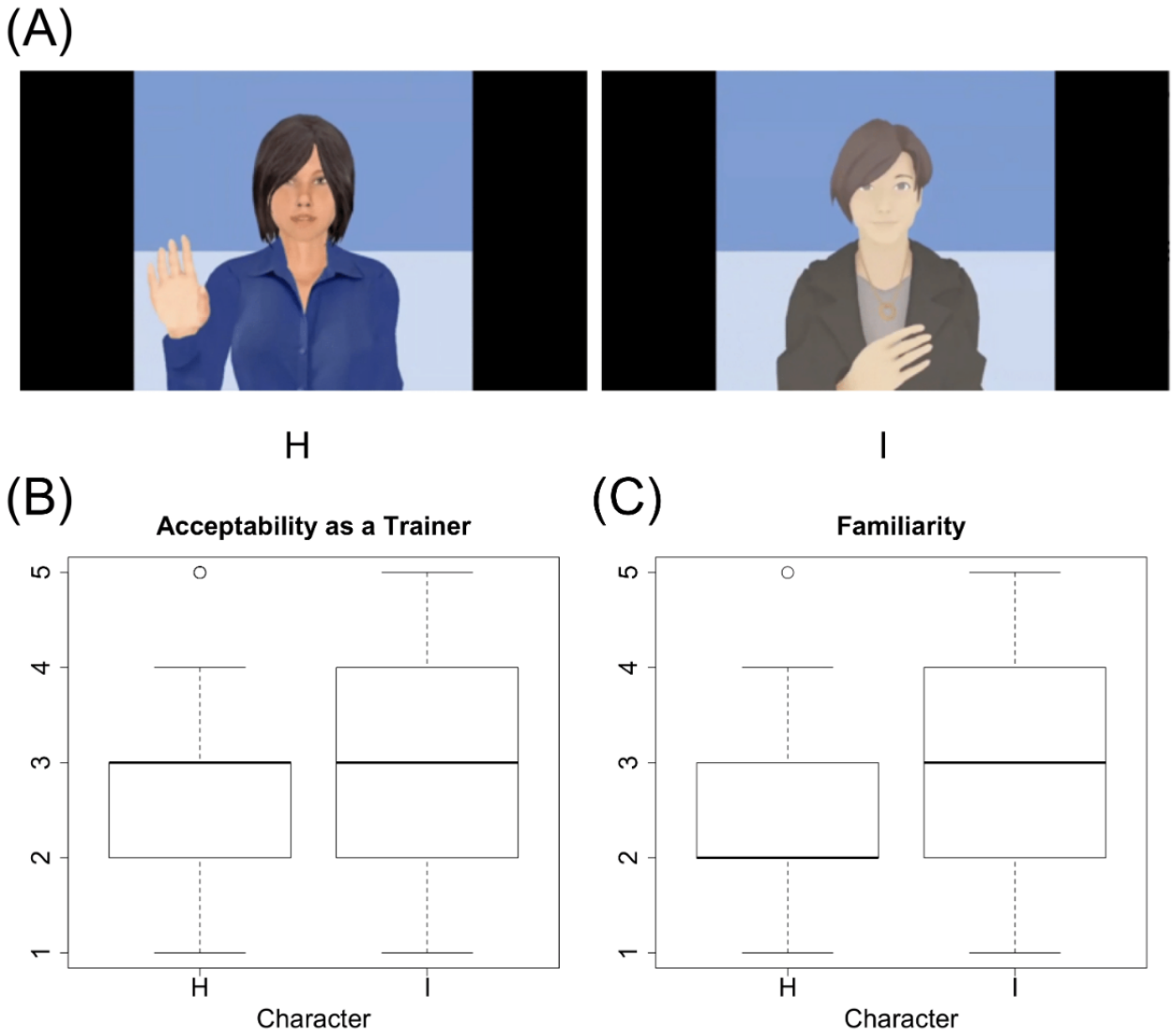
All characters faced the front with no emotional expression. Characters A and B (female) and C and D (male) were designed with a consistent age, with only the degree of their realism and gender changing. Character E was an inanimate object created for use with children. Character F was a nonhuman animal (a dog), also for use with children. For character G, we created a

realistic 3D model similar in appearance to characters A and B and took a screen capture from the front. Character H was the default agent provided by the Greta platform (developed by Pelachaud et al) [29], which is an embodied conversation agent that can be created with the Autodesk character generator (Autodesk Inc.) [30]. Character H was intended for use mainly with French- and English-speaking users. Character F was designed for Japanese female users. In the current study of automated social skills training, character I was selected as the virtual character [10]. The representation of characters H and I was created by taking a screen capture from the front.

The sentence "Hello, let's practice communication together" was embedded in the image with both a male and a female voice. The utterance was 5 seconds in length and spoken by Google Text-to-Speech. Characters E and F were created with higher-pitched voices than those used for normal female speech synthesis, to mimic children's voices.

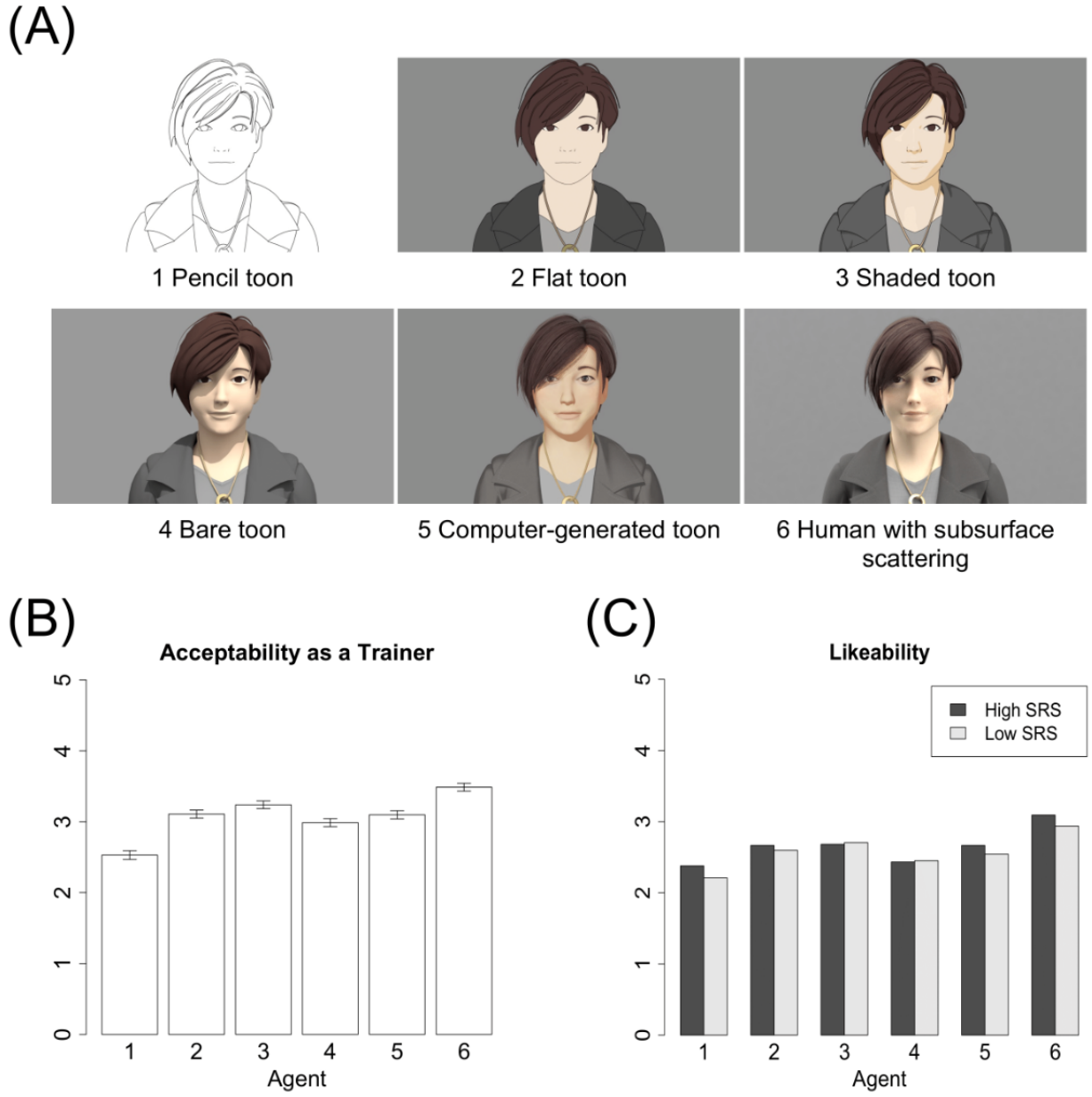
Since 3D models were available for characters H and I, we were also able to create videos for them in Greta (Figure 2). Movements and gestures were added, such as the character raising its hands or putting its hands on its chest, synchronized to the speech content. These same behaviors and synchronization for characters H and I were also generated in Japanese, with an utterance length of 8 seconds. The speech synthesis used the voice of the character "Yuki" in CereProc (CereProc Ltd.).

Figure 2. Screen captures of videos of two of the virtual characters (A) and representative measures collected from data set 3 (B, C).



We further analyzed the effect of realism by designing additional virtual agents, also with the aid of a design company specializing in Japanese animation. These agents were designed using the Maya tool (Autodesk Inc.). We prepared 6 levels of realism, following a previously reported method [23]. The degrees of realism were as follows: (1) pencil toon, (2) flat toon, (3) shaded toon, (4) bare toon, (5) computer-generated toon, and (6) human (with subsurface scattering), as shown in Figure 3. The same behavior was generated for these 6 agents, with Japanese speech synthesis and lip-synching using the same words as described above. All of these images and movies are available upon request to the first author.

Figure 3. Screen captures of the 6 virtual agents (A); acceptability as a trainer in data set 4 (error bars represent SE) (B); and the evaluation of likeability by high and low SRS score groups (C).



Participants

For data collection, we recruited participants from a crowdsourcing service (Crowdworks). The recruitment notice asked for participants 18 years of age or older with Japanese nationality. In order to divide the task among the participants, data were collected in 4 separate data sets with different participants. Data set 1 had 305 participants (with a male to female ratio of 148 to 157), data set 2 had 301 participants (with a male to female ratio of 131 to 170), data set 3 had 302 participants (with a male to female ratio of 145 to 157), and data set 4 had 305 participants (with a male to female ratio of 145 to 160). All data sets can be found in multimedia appendix. Data set 1 was used to investigate image acceptability, likeability, familiarity, likeability of certain elements (ie, eyes, face, hair, voice, and perceived age), autistic traits, and alexithymia. Data set 2 was used to investigate realism, trustworthiness, and eeriness of the agents. Data set 3 was used to investigate the videos with characters H and I. Data set 4 was used to investigate the realism of the movies, as well as autistic traits. For the validation to have a sufficient sample size, we collected a larger sample size for each data set compared to previous works, which have recruited around 40 to 70 participants from regional communities [20,21] or have used Amazon's Mechanical Turk platform [24]. In this study, we also performed a grouped analysis using 45 years as the threshold for high and low age groups (high age: n=84; low age: n=21).

Autistic Traits

In data set 1 and data set 4, we used the adult version of the Social Responsiveness Scale-2 (SRS) [31] to assess autistic traits. This measures how many autistic traits an individual shows and can be used across the general population, not only with people who are suspected of having ASD. In data set 1, we measured the Toronto Alexithymia Scale-20 (TAS) [32] to

assess alexithymia. In both cases, we calculated the total score. We did not calculate subscales in this study. In data set 1, the 2 questionnaires had a Spearman correlation coefficient of 0.67 (P<.001), which indicates a high correlation between autistic traits and alexithymia. We are currently planning a future analysis that will use SRS as a measure of autistic traits. In this study, we used a cutoff value of 81 points [33] as the threshold for high and low SRS score (subjects with a high SRS score: n=113; low SRS score: n=192). We also measured SRS scores in data set 4 and also set a threshold for high and low SRS scores in that data set (high: n=129, low: n=177).

Measures

Questionnaire items and scales were prepared with reference to studies by Esposito et al and Ring et al [20,21,24,25]. The questionnaire items measured the acceptability of the agent as a trainer and as a listener; its realism, familiarity, trustworthiness, and eeriness; the likeability of its face, eyes, hair, perceived age, and voice; and its overall impression. Each question was answered through a Google Form. In data set 1, each question item was answered after completing the SRS and TAS. In data set 3, in addition to the above, we added the likeability of the clothes the agent wore, because the video included the entire upper body of the virtual agent. We asked the participants to read a description of the concept of social skills training (in particular, the function of a virtual agent to train the user's social communication skills and also listen to the user). We performed a preliminary test with a few adults to check whether the participants understood the social skills training, and we wrote instructions. Participants first looked at a set of all the images (Figure 1) to get an impression of all the virtual agents, and they then watched the individual virtual agents and answered each question. The questions were evaluated with a 5-point Likert scale (from 1, "I don't think so at all," to 5 "I think so very much"). Spearman ρ was calculated to determine the relationship between the questionnaire items.

R (R Foundation for Statistical Computing) was used for the analysis. Since normality could not be confirmed in the ratings of the questions by the Kolmogorov-Smirnov test, the Kruskal-Wallis test was used to examine the differences between the virtual agents. In the analysis for each group of gender, age, and SRS, we calculated the effect size (r). We report the top 3 combinations of r from all combinations of virtual agents and questionnaire items. Furthermore, we performed the Wilcoxon signed-rank test to compare pairs of factors.

Ethical Considerations

This was an anonymous study in which the participants enrolled themselves by registering through Crowdworks and agreeing to participate in the study. Since participation was anonymized, the study was exempt from registration with our institutional review board.

Results

In reporting the results, we did not report all measures, in order to focus on significant findings. The following is a summary of the experimental results.

First, the differences in ratings between the virtual characters. The Kruskal-Wallis test confirmed that there were significant differences between the virtual characters in all measures (P<.001). Regarding realism, the distribution was as expected in the original design: character A was more realistic than character B and character G was the most realistic. The most preferred virtual character among the participants was character B, averaging 3.29 (SD 1.0) (Figure 1). Character B was also highly evaluated in other questionnaire items. We also found that the male characters, C and D, and the nonhuman characters, E and F, had lower likeability than character B, and that character H had less likeability and less familiarity.

Next, the correlations between questionnaire items. Figure 4 shows the correlation matrix. There was a high correlation between face and preference (ρ=0.78, P<.001). There was also a high correlation between acceptance as a trainer and acceptance as a listener (ρ=0.80, P<.001). On the other hand, although a significant difference was confirmed regarding voice preference and other questionnaire items, the correlation coefficient was relatively low.

Table 1 lists the top 3 combinations of virtual agents and questionnaire items that had the highest effect size (r) for gender, age, and SRS score. All cases with a statistically significant difference are listed in Multimedia Appendix 1. Male subjects evaluated character G's face, overall likeability, and acceptability as a trainer more highly than did female subjects. The higher age group evaluated character I's eyes and face more highly than did the lower age group. The high SRS score group evaluated the likeability of character G's eyes and hair more highly than did the low SRS score group.

Figure 2 shows a comparison of the videos of characters H and I, indicating that acceptability and familiarity were significantly greater for character I than H (all P<.001).

Figure 5 shows the overall rating for realism for the agents shown in Figure 3. The Kruskal-Wallis test confirmed that the virtual agents differed significantly in realism (P<.001) and confirmed our design assumption that agents 1 through 6 would have increasingly greater realism. Figure 3 shows the acceptability as a trainer and likeability of the virtual agents. The Kruskal-Wallis test confirmed that the virtual agents differed significantly in all measures (P<.001). We found a small difference between the high and low SRS score groups in their evaluation of likeability (Figure 3 lower right), but the Wilcoxon rank-sum test showed no significant difference (for character 1, P=.13 and for character 6, P=.25) and a small effect size.

Figure 4. Correlation matrix of measures.

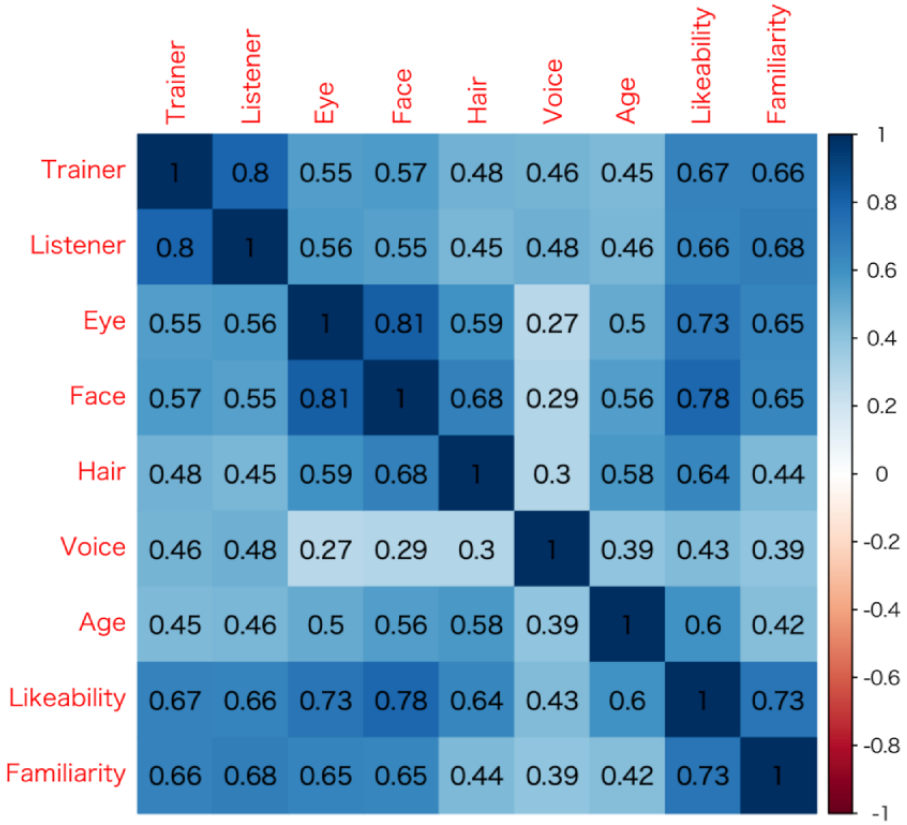
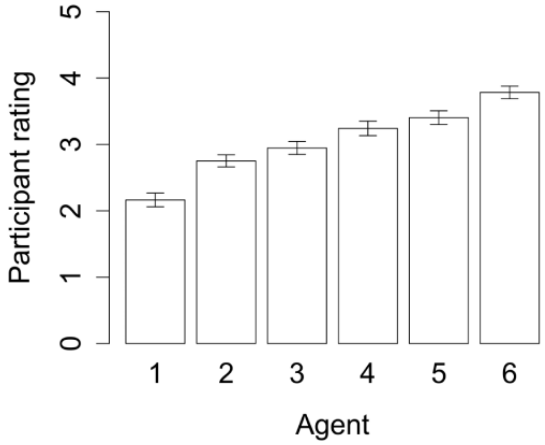


Table 1. Relationship between questionnaire items and gender, age, and SRS score.

User characteristic	Questionnaire item	*r* (*P* value)	Trend
Gender			
Character G	Face	0.29 (<.001)	Male > female
Character G	Likeability	0.25 (<.001)	Male > female
Character G	Trainer	0.25 (<.001)	Male > female
Age			
Character I	Eyes	0.21 (<.001)	High > low
Character I	Face	0.19 (<.001)	High > low
Character A	Listener	0.17 (.003)	High < low
SRS score			
Character G	Eyes	0.19 (.001)	High > low
Character G	Hair	0.18 (.002)	High > low
Character G	Face	0.16 (.009)	High > low

Figure 5. Realism measures collected from data set 4. Error bars represent SE.



Discussion

The objective of this study was to examine virtual agent visual design for automated social skills training, the relationship between acceptability and other measures, and the relationship between likeability and individual user characteristics. We also investigated the acceptability and likeability of the virtual agents, as well as various other measures. We were able to confirm that the virtual agents had different ratings. First, we found that the realism of the virtual agent design could be controlled through the selection of characters A, B, or G. We found that character B, originally designed as an anime-like teenage female character, was the most likable (Figure 1). Since Japanese people are rather accustomed to watching anime-like videos, familiarity with such characters is high. The anime art form, having originated in Japan in the early 1900s, is a uniquely stylized form of 2D and 3D illustration [34]. Such a female anime-like character was also integrated and familiarized in our previous research on automated social skills training [8]. On the other hand, other virtual characters, such as the inanimate object (character E) or the animal (character F), as well as characters G and H, were less accepted and were not preferred.

We found significant correlations between questionnaire items (P<.001) and a high correlation between face and preference. This face factor influenced the development of the automated social skills training. There was also a high correlation between acceptance as a trainer and acceptance as a listener (Figure 4). In this case, we could not confirm the difference between the role as trainer and that as listener, because no continuous interactive dialogue was available. When the roles of virtual characters are more carefully chosen in the future, we assume that an investigation of this issue will also be necessary. Since the same voice was used for each virtual character, the correlation coefficient was relatively low. Therefore, we should explore the effect of voice using a variety of speech synthesizers in the future.

We also found very similar tendencies in video versions of the training agents. However, in terms of familiarity, we confirmed that the rating for the video version of character H was higher than its image version due to the addition of naturalistic movement. Regarding the videos shown in Figure 2, acceptability, and familiarity were significantly greater for character I than H. This shows that Japanese users preferred the anime-like character I over the original Greta character H.

We also found that realism, as shown in Figure 5, was associated with acceptability and likeability (Figure 3), a finding that is similar to that reported by McDonnell et al [23]. This may be related to the uncanny valley effect [35,36] and represent an intermediary between the responses to characters 3 (shaded toon) and 4 (bare toon). Although the most highly evaluated agent was character 6, the human with subsurface scattering, this sort of agent may need high-quality 3D modeling for its appearance and movement to be natural enough for use in automated social skills training. Thus, the second-ranked character, character 3 (shaded toon), may be the most promising for a realistic virtual agent for automated social skills training.

We found that the female virtual character, character G, was rated as more preferred by male participants. In addition, since we confirmed that character B was also significantly highly rated by male participants, it appears that the male participants rated female virtual characters as more preferable. Character B, originally designed as an anime-like teenage female character, was judged the most likable by all participants. We found that character I was preferred by older participants. Since character I was designed to appear relatively older (and was originally designed for participants in their 40s), it seems that the older group rated characters closer to their own age as more trustworthy. Therefore, when developing automated social skills training for older users, character I might be the most appropriate type of visual design. In this paper, one of our goals was to analyze the effect of autistic traits. We found that autistic traits were strongly associated with alexithymia (Spearman ρ=0.67). Thus, we focused only on SRS score to measure autistic traits. Our results showed that people with high autistic traits had a preference for realistic agents. We also confirmed that the group with high autistic traits gave a high rating to characters G and H in data set 1. This is a similar finding to previous work [18]. However, we did not find a difference in the case of data set 4.

Further investigation is needed to examine altered cognition in autism and its effects in order to conduct a comparison of virtual agents and real human agents. Although the target population of this study was adults 18 years or older, children with ASD may prefer nonhuman virtual agents, such as trains [5]. We must consider the effects of virtual agents in younger users. In future

work, we hope to examine the effect of cultural differences, younger age, and virtual agent facial expressions on acceptability. These features could be used as variables of interest. In addition, this study did not confirm whether crowdsourced workers have sufficient knowledge of social skills training. Consequently, we need to investigate the effects of integrating design into an interactive social skills training dialogue system [8,9].

Conclusions

In this study, we prepared various new virtual agent visual designs for social skills training and evaluated the designs based on multiple questionnaire items that assessed likeability, acceptability, realism, familiarity, and trustworthiness, among other factors, in a study sample of 1218 crowdsourced evaluators. We tested differences in preferences for virtual agent visual designs based on the gender, age, and autistic traits of the participants, in order to create personalized virtual agents. We found that our participants preferred, perhaps through familiarity, anime-like characters, likely because Japanese people are rather accustomed to watching anime-like videos. Our conclusion for implementing an optimal virtual agent for use with Japanese users is generally to design a female anime-type agent (especially a toon-shaded type), which has been shown to be favored and acceptable. We also found that preferences for virtual agent visual design differed according to user gender, age, and autistic traits. For example, we confirmed that users with high autistic traits showed a high preference for virtual agents with a realistic appearance.

Acknowledgments

This work was funded by Japan Science and Technology Agency CREST (grant JPMJCR19M5) and Japan Society for the Promotion of Science KAKENHI (grants JP17H06101 and JP18K11437).

Conflicts of Interest

None declared.

Multimedia Appendix 1
All cases with a statistical difference.
[DOCX File , 14 KB - humanfactors_v9i1e35358_app1.docx]

Multimedia Appendix 2
Data set 1 to 4.
[ZIP File (Zip Archive), 58 KB - humanfactors_v9i1e35358_app2.zip]

References

1. Bellack AS, Meuser KT, Gingerich S, Agresta J. Social Skills Training for Schizophrenia: A Step-by-Step Guide. New York, NY: Guilford Press; 1997:637-638.
2. Liberman RP, Mueser KT, Wallace CJ. Social skills training for schizophrenic individuals at risk for relapse. Am J Psychiatry 1986 Apr;143(4):523-526. [doi: 10.1176/ajp.143.4.523] [Medline: 2869704]
3. Skinner BF. Science And Human Behavior. New York: Macmillan; 1953.
4. Frith U, Happé F. Autism spectrum disorder. Curr Biol 2005 Oct 11;15(19):R786-R790 [FREE Full text] [doi: 10.1016/j.cub.2005.09.033] [Medline: 16213805]
5. Golan O, Baron-Cohen S. Systemizing empathy: teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. Dev Psychopathol 2006;18(2):591-617. [doi: 10.1017/S0954579406060305] [Medline: 16600069]
6. Tanaka H, Negoro H, Iwasaka H, Nakamura S. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. PLoS One 2017;12(8):e0182151 [FREE Full text] [doi: 10.1371/journal.pone.0182151] [Medline: 28796781]
7. Poyade M, Morris G, Taylor I, Portela V. Using mobile virtual reality to empower people with hidden disabilities to overcome their barriers. 2017 Presented at: The 19th ACM International Conference on Multimodal Interaction; Nov 13-17, 2017; Glasgow, Scotland p. 504-505. [doi: 10.1145/3136755.3143025]
8. Tanaka H, Negoro H, Iwasaka H, Nakamura S. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. PLoS One 2017 Aug 10;12(8):e0182151 [FREE Full text] [doi: 10.1371/journal.pone.0182151] [Medline: 28796781]
9. Tanaka H, Iwasaka H, Negoro H, Nakamura S. Analysis of conversational listening skills toward agent-based social skills training. J Multimodal User Interfaces 2019 Oct 16;14(1):73-82. [doi: 10.1007/s12193-019-00313-y]
10. Tanaka H, Iwasaka H, Matsuda Y, Okazaki K, Nakamura S. Analyzing Self-Efficacy and Summary Feedback in Automated Social Skills Training. IEEE Open J. Eng. Med. Biol 2021;2:65-70. [doi: 10.1109/ojemb.2021.3075567]

11. Ali MR, Rasazi Z, Mamun AA, Langevin R, Rawassizadeh R, Schubert LK, et al. A Virtual Conversational Agent for Teens with Autism: Experimental Results and Design Lessons. CoRR 2018:E. [doi: 10.1145/3383652.3423900]
12. Hoque M, Courgeon M, Martin JC, Mutlu B, Picard RW. MACH: my automated conversation coach, UbiComp, pp. 2013 Presented at: the 2013 ACM international joint conference on pervasive and ubiquitous computing (UbiComp '13); Sep 8-12, 2013; Zurich, Switzerland. [doi: 10.1145/2493432.2493502]
13. Tudor Car L, Dhinagaran DA, Kyaw BM, Kowatsch T, Joty S, Theng Y, et al. Conversational Agents in Health Care: Scoping Review and Conceptual Analysis. J Med Internet Res 2020 Aug 07;22(8):e17158 [FREE Full text] [doi: 10.2196/17158] [Medline: 32763886]
14. Milne-Ives M, de Cock C, Lim E, Shehadeh MH, de Pennington N, Mole G, et al. The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review. J Med Internet Res 2020 Oct 22;22(10):e20346 [FREE Full text] [doi: 10.2196/20346] [Medline: 33090118]
15. Tanaka H, Sakriani S, Neubig G, Toda T, Negoro H, Iwasaka H, et al. Teaching Social Communication Skills Through Human-Agent Interaction. ACM Trans. Interact. Intell. Syst 2016 Aug 03;6(2):1-26. [doi: 10.1145/2937757]
16. Shidara K, Tanaka H, Adachi H, Kanayama D, Sakagami Y, Kudo T, et al. Automatic Thoughts and Facial Expressions in Cognitive Restructuring With Virtual Agents. Front. Comput. Sci 2022 Feb 2;4. [doi: 10.3389/fcomp.2022.762424]
17. Ardito RB, Rabellino D. Therapeutic alliance and outcome of psychotherapy: historical excursus, measurements, and prospects for research. Front Psychol 2011;2:270 [FREE Full text] [doi: 10.3389/fpsyg.2011.00270] [Medline: 22028698]
18. Kumazaki H, Warren Z, Muramatsu T, Yoshikawa Y, Matsumoto Y, Miyao M, et al. A pilot study for robot appearance preferences among high-functioning individuals with autism spectrum disorder: Implications for therapeutic use. PLoS One 2017 Oct 13;12(10):e0186581 [FREE Full text] [doi: 10.1371/journal.pone.0186581] [Medline: 29028837]
19. Robins B, Dautenhahn K, Te Boekhorst R, Billard A. Robots as assistive technology - does appearance matter? 2004 Presented at: International Workshop on Robot and Human Interactive Communication, 13th IEEE ROMAN; Sep 22-24, 2004; Kurashiki, Japan p. 277-282. [doi: 10.1109/roman.2004.1374773]
20. Esposito A, Amorese T, Cucinello M, Esposito AM, Troncone A, Ines Torres M, et al. Seniors' Acceptance of Virtual Humanoid Agents. Italian Forum of Ambient Assisted Living 2018:429. [doi: 10.48550/arXiv.2105.00506]
21. Esposito A, Amorese T, Cuciniello M, Pica I, Riviello MT, Troncone A, et al. Elders Prefer Female Robots with a High Degree of Human Likeness. 2017 Presented at: IEEE 23rd International Symposium on Consumer Technologies; Jun 19-21, 2017; Ancona, Italy p. 243-246. [doi: 10.1109/isce.2019.8900983]
22. Terada K, Jing L, Yamada S. Effects of Agent Appearance on Customer Buying Motivations on Online Shopping Sites. 2015 Presented at: the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems; 2015; Seoul, Korea p. 929-934. [doi: 10.1145/2702613.2732798]
23. McDonnell R, Breidt M, Bülthoff HH. Render me real?: investigating the effect of render style on the perception of animated virtual humans. ACM Trans. Graph 2012 Aug 05;31(4):1-11. [doi: 10.1145/2185520.2185587]
24. Ring L, Utami D, Bickmore T. The Right Agent for the Job? The Effects of Agent Visual Appearance on Task Domain. 2014 Presented at: International Conference on Intelligent Virtual Agents; Aug 27-29, 2014; Boston, MA p. 374-384. [doi: 10.1007/978-3-319-09767-1_49]
25. Esposito A, Amorese T, Cuciniello M, Riviello MT, Esposito AM, Troncone A, et al. The Dependability of Voice on Elders' Acceptance of Humanoid Agents. 2019 Presented at: Interspeech 2019; Sep 15-19, 2019; Graz, Austria p. 31. [doi: 10.21437/interspeech.2019-1734]
26. Baylor AL, Yanghee K. Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role. 2004 Presented at: Intelligent Tutoring Systems, 7th International Conference, ITS 2004; Aug 30-Sep 3, 2004; Maceiò, Alagoas, Brazil p. 592. [doi: 10.1007/978-3-540-30139-4_56]
27. Troncone A, Amorese T, Cuciniello M, Saturno R, Pugliese L, Cordasco G, et al. Advanced Assistive Technologies for Elderly People: A Psychological Perspective on Seniors' Needs and Preferences (part A). ACTA POLYTECH HUNG 2020;17(2):163-189. [doi: 10.12700/aph.17.2.2020.2.10]
28. Tanaka H, Nakamura S. Virtual Agent Design for Social Skills Training Considering Autistic Traits. 2020 Presented at: 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Jul 20-24, 2020; Montreal, Canada p. 4959-4962. [doi: 10.1109/embc46164.2021.9630741]
29. Poggi C, Pelachaud C, De Rosis F, Carofiglio V, De Carolis B. Greta. A Believable Embodied Conversational Agent. In: Multimodal Intelligent Information Presentation. Dordrecht: Springer; 2005:3-25.
30. Autodesk Character Generator. URL: https://charactergenerator.autodesk.com/ [accessed 2022-03-21]
31. Constantino JN. Social Responsiveness Scale - Second Edition (SRS-2), WPS. In: Social Responsiveness Scale - Second Edition (SRS-2). Los Angeles, CA: Western Psychological Services; 2012.
32. Bagby R, Parker JD, Taylor GJ. The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. Journal of Psychosomatic Research 1994 Jan;38(1):23-32. [doi: 10.1016/0022-3999(94)90005-1]
33. Bezemer ML, Blijd-Hoogewys EMA, Meek-Heekelaar M. The Predictive Value of the AQ and the SRS-A in the Diagnosis of ASD in Adults in Clinical Practice. J Autism Dev Disord 2021 Jul 01;51(7):2402-2415 [FREE Full text] [doi: 10.1007/s10803-020-04699-7] [Medline: 33001348]
34. Anime Art Museum. URL: http://animeartmuseum.org/whatisanimeart/ [accessed 2022-03-21]

35. Koschate M, Potter R, Bremner P, Levine M. Overcoming the uncanny valley: Displays of emotions reduce the uncanniness of humanlike robots. 2016 Presented at: 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI); Mar 7-10, 2016; Christchurch, New Zealand. [doi: 10.1109/hri.2016.7451773]
36. Mori M. The uncanny valley. Energy 7 1970:33. [doi: 10.5749/j.ctvtv937f.7]

Edited by A Kushniruk; submitted 02.12.21; peer-reviewed by S Okada, M Chollet, B Chaudhry; comments to author 20.01.22; revised version received 30.01.22; accepted 21.02.22; published 29.03.22.

Please cite as:
Tanaka H, Nakamura S
The Acceptability of Virtual Characters as Social Skills Trainers: Usability Study
JMIR Hum Factors 2022;9(1):e35358
URL: https://humanfactors.jmir.org/2022/1/e35358
doi:10.2196/35358
PMID:35348468

Letter to the Editor

Ensuring Interrater Reliability When Evaluating Voice Assistants. Comment on “Evaluating Voice Assistants’ Responses to COVID-19 Vaccination in Portuguese: Quality Assessment”

Rujittika Mungmunpuntipantip[1*], PhD; Viroj Wiwanitkit[2*], MD

[1]Private Academic Consultant, Bangkok, Thailand

[2]Department of Community Medicine, DY Patil University, Pune, Thailand

[*]all authors contributed equally

Corresponding Author:
Rujittika Mungmunpuntipantip, PhD
Private Academic Consultant
26 Bangkok 111
Bangkok, 1-101132
Thailand
Phone: 66 2388282822
Email: rujittika@gmail.com

Related Article:

Comment on: http://www.jmir.org/2022/1/e34674/

(JMIR Hum Factors 2022;9(1):e36610) doi:10.2196/36610

KEYWORDS

voice assistant; natural user interface; Portuguese language; COVID-19; vaccine

We would like to share our ideas on the paper “Evaluating Voice Assistants’ Responses to COVID-19 Vaccination in Portuguese: Quality Assessment” [1]. Seródio Figueiredo et al [1] concluded, “Under the urgent context of COVID-19 vaccination, this work can help to understand how VAs must be improved to be more useful to the society and how careful people must be when considering VAs as a source of health information.” We agree that voice assistants (VAs) could be useful in managing COVID-19 mass immunization campaigns. Current reports can provide an overview of existing technologies and the mission of various VA suppliers in order to meet current information distribution requirements. However, as previously said, the question of system functioning remains critical. Findings on VAs still have variability and need to be harmonized [2]. In addition, effective governance is required for transparency and usefulness in information partnerships [2].

Seródio Figueiredo et al [1] assessed VAs via 2 evaluators. A set of questions was used. The focus was on agreement between the 2 evaluators. Interevaluator variability should be assessed and presented as well. In general, there should be 3 evaluators to aid the final decision. Additionally, the scoring system of this study might be easily influenced by bias. Essentially, any questionnaire-based study requires a reliability test of the questionnaire. Content validity, face validity, and criterion-related validity tests are required. A standard method, as presented by Bolarinwa [3], should be used. A statistical analysis of the questionnaire’s reliability must be presented in addition to a general description. Without proven measures of questionnaire reliability and evaluator variability, the use of this tool in this study might be questionable.

Conflicts of Interest

None declared.

Editorial Notice

The corresponding author of “Evaluating Voice Assistants’ Responses to COVID-19 Vaccination in Portuguese: Quality Assessment” declined to respond to this letter.

References

1. Seródio Figueiredo CM, de Melo T, Goes R. Evaluating Voice Assistants' Responses to COVID-19 Vaccination in Portuguese: Quality Assessment. JMIR Hum Factors 2022 Mar;9(1):e34674 [FREE Full text] [doi: 10.2196/34674] [Medline: 35041617]
2. Alagha EC, Helbing RR. Evaluating the quality of voice assistants' responses to consumer health questions about vaccines: an exploratory comparison of Alexa, Google Assistant and Siri. BMJ Health Care Inform 2019 Nov;26(1) [FREE Full text] [doi: 10.1136/bmjhci-2019-100075] [Medline: 31767629]
3. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. Niger Postgrad Med J 2015;22(4):195-201 [FREE Full text] [doi: 10.4103/1117-1936.173959] [Medline: 26776330]

Abbreviations

VA: voice assistant

Edited by T Leung; submitted 18.01.22; this is a non–peer-reviewed article;accepted 25.01.22; published 21.03.22.

Please cite as:
Mungmunpuntipantip R, Wiwanitkit V
Ensuring Interrater Reliability When Evaluating Voice Assistants. Comment on "Evaluating Voice Assistants' Responses to COVID-19 Vaccination in Portuguese: Quality Assessment"
JMIR Hum Factors 2022;9(1):e36610
URL: https://humanfactors.jmir.org/2022/1/e36610
doi:10.2196/36610
PMID:35312626

Original Paper

Trends in Remote Health Care Consumption in Sweden: Comparison Before and During the First Wave of the COVID-19 Pandemic

Veronica Milos Nymberg[1*], PhD; Lina Maria Ellegård[2,3*], PhD; Gustav Kjellsson[4*], PhD; Moa Wolff[1*], PhD; Beata Borgström Bolmsjö[1*], PhD; Thorne Wallman[5,6*], PhD; Susanna Calling[1*], PhD

[1]Center for Primary Health Care Research, Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden

[2]Department of Economics, Lund University, Lund, Sweden

[3]Department of Business Administration, Kristianstad University, Kristianstad, Sweden

[4]Department of Economics, University of Gothenburg, Gothenburg, Sweden

[5]Family Medicine & Preventive Medicine Section, Public Health & Caring Sciences, Uppsala University, Uppsala, Sweden

[6]Center for Clinical Research Sörmland, Uppsala university, Eskilstuna, Sweden

[*]all authors contributed equally

Corresponding Author:
Veronica Milos Nymberg, PhD
Center for Primary Health Care Research
Department of Clinical Sciences Malmö
Lund University
Building 28, Floor 11
Jan Waldenströms Street 35
Malmö, 20502
Sweden
Phone: 46 767700240
Email: veronica.milos_nymberg@med.lu.se

Abstract

Background: Remote assessment of respiratory tract infections (RTIs) has been a controversial topic during the fast development of private telemedicine providers in Swedish primary health care. The possibility to unburden the traditional care has been put against a questionable quality of care as well as risks of increased utilization and costs. The COVID-19 pandemic has contributed to a changed management of patient care to decrease viral spread, with an expected shift in contact types from in-person to remote ones.

Objective: The main aim of this study was to compare health care consumption and type of contacts (in-person or remote) for RTIs before and during the COVID-19 pandemic. The second aim was to study whether the number of follow-up contacts after an index contact for RTIs changed during the study period, and whether the number of follow-up contacts differed if the index contact was in-person or remote. A third aim was to study whether the pattern of follow-up contacts differed depending on whether the index contact was with a traditional or a private telemedicine provider.

Methods: The study design was an observational retrospective analysis with a description of all index contacts and follow-up contacts with physicians in primary care and emergency rooms in a Swedish region (Skåne) for RTIs including patients of all ages and comparison for the same periods in 2018, 2019, and 2020.

Results: Compared with 2018 and 2019, there were fewer index contacts for RTIs per 1000 inhabitants in 2020. By contrast, the number of follow-up contacts, both per 1000 inhabitants and per index contact, was higher in 2020. The composition of both index and follow-up contacts changed as the share of remote contacts, in particular for traditional care providers, increased.

Conclusions: During the COVID-19 pandemic in 2020, fewer index contacts for RTIs but more follow-up contacts were conducted, compared with 2018-2019. The share of both index and follow-up contacts that were conducted remotely increased. Further studies are needed to study the reasons behind the increase in remote contacts, and if it will last after the pandemic, and more clinical guidelines for remote assessments of RTI are warranted.

(JMIR Hum Factors 2022;9(1):e33034) doi:10.2196/33034

KEYWORDS

remote health care; telemedicine; primary health care; respiratory tract infections; COVID-19

Introduction

Respiratory tract infections (RTIs) are one of the most common reasons for contacts in Swedish primary care [1]. The outbreak of the COVID-19 pandemic has led to a higher threshold to assess uncomplicated RTIs in primary care using in-person contacts. Thus, the pandemic has catalyzed the development, implementation, and use of remote contacts with primary care providers, including traditional telephone contacts as well as digital contacts (email, chat, video consultations) [2]. Private telemedicine providers offering exclusively video consultations or chat on-demand have been established both in Sweden and worldwide during past years [3,4], and the number of contacts with private telemedicine providers is increasing rapidly [4]. The traditional primary care sector in Sweden has also implemented different telemedicine platforms as a complement to traditional care, including mostly chat or email. Although patients and health care staff report satisfaction with remote contacts [5,6], there is a lack of evidence about risks, benefits, and cost efficiency regarding assessment of symptoms through digital contacts compared with traditional physical contacts in primary care [7-9]. For example, diagnostic difficulties and an increased number of follow-up contacts have been described after management of RTIs by digital contacts in the United States [10]. However, a recent Swedish study found no increase in follow-up rates or antibiotic use for virtual care compared with emergency care or primary care for low-acuity urgent conditions [11]. Previous studies have described the expansion of telemedicine in Sweden before 2019 [12]. In line with international evidence [8], the availability of such services has been associated with a net increase in health care utilization [13].

The COVID-19 pandemic has led to a fast shift in the modality by which patients and health care communicate, from in-person to remote contacts (primarily to minimize the SARS-CoV-2 spread), despite insufficient guidelines and conflicting evidence regarding its impact on health-related outcomes [14]. A basic search in PubMed (MEDLINE) using keywords of the study aim identified a knowledge gap, with no prior studies about how this shift in communication has affected the search pattern for RTIs, and how the composition of follow-up contacts differs if the index contact is physical or digital.

The first aim of this study was to analyze the change in number and percentage of in-person contacts and remote contacts, respectively, for RTI between January-June 2018/19 and January-June 2020. The second aim was to study whether the number of follow-up contacts after an index contact for RTI changed during the study period, and whether the type and number of follow-up contacts differed between index in-person contacts and index remote contacts. A third aim was to study whether the pattern of follow-up contacts differed depending on whether the index contact was with a traditional care provider or a private telemedicine provider.

Methods

Study Design

The study design was an observational retrospective analysis describing physician contacts in primary care and at hospital emergency rooms of patients with an index contact for RTIs in January-June 2020, 2019, and 2018. The first 6 months of 2020 were chosen to study the development of RTI-related contacts during the start of the COVID-19 pandemic. The monthly development of physician contacts in 2018 and 2019 was used to illustrate seasonal patterns in the absence of a pandemic. The study was set in the Swedish region Skåne, which is the third largest region (1.4 million inhabitants), with a wide geographical variation including large, middle-sized, and small cities as well as rural areas. The region was relatively mildly hit by the first wave of the pandemic. In the first half of 2020, the number of laboratory-confirmed COVID-19 cases per 100,000 inhabitants was 219; 110 patients with COVID-19 received intensive care and 248 inhabitants died with a COVID-19 diagnosis. This may be compared with one of the worst hit regions in Sweden—Stockholm (2.4 million inhabitants), which recorded 798 confirmed cases per 100,000 inhabitants, 894 intensive care patients, and 2,331 deaths in the same period [15].

Study Population and Data

The study population consisted of all individuals with a registered address in Skåne (Region Skåne) on December 31 in 2017, 2018, or 2019 (according to the Swedish population register held by Statistics Sweden). For this population, data on in-person or remote contacts with care providers located in Skåne (public or contracting with the region) were collected from the regional health authority's care register "Region Skånes Vårddatabas" (RSVD). The data included information on date of contact, type of contact (in-person/remote), and up to 8 diagnoses for all care contacts in the period August 2017 to July 2020. Data for the same population located in Skåne on contacts with private telemedicine providers, which are formally located in other regions (Region Sörmland and Region Jönköping) and therefore report to care registers in these regions, were sourced from the health authorities in those regions and a register of extra-regional care contacts of inhabitants in Skåne. Multimedia Appendix 1 presents more details on data sources for the different types of contacts and information about missing data on registered diagnoses from the various sources.

In the analysis, we distinguished between physician contacts that were in-person or remote (ie, consultations by telephone, video, or asynchronous chats). We further distinguished between remote contacts with traditional providers (defined as all primary health care centers and hospital emergency rooms) and remote contacts with pure on-demand telemedicine providers. Such telemedicine services were offered by the private companies Kry, Capio Go, Min Doktor, Doktor.se, Doktor 24, Medicoo, Accumbo. During the study period, traditional care providers

mainly offered remote contacts via phone and to some extent by asynchronous chats, whereas private on-demand telemedicine providers primarily offered asynchronous chats or video calls. We linked the data from different registers using pseudo-anonymous individual identifiers provided by Statistics Sweden. The linked data set included all contacts with traditional and private telemedicine providers made by the study population.

Outcome Variables

Primary outcome variables were *index contacts* and *follow-up contacts with a physician* for patients with a registered RTI-relevant diagnosis (Table 1). Definition of the variables are presented in Table 2.

The study included index contacts occurring during any of the following periods: January 1, 2018-June 30, 2018; January 1, 2019-June 30, 2019; and January 1, 2020-June 30, 2020. We studied the total number of index contacts per 1000 inhabitants, and the total number of follow-up contacts per 1000 inhabitants and per index contact. In addition, we reported the number of unique patients hospitalized with an RTI diagnosis according to Table 1 (per 1000 inhabitants).

Table 1. ICD-10[a] diagnosis codes of relevant diagnoses [16].

Diagnosis code ICD-10	Diagnosis group
J00-J06	Acute upper respiratory infections
J10-J18	Influenza and pneumonia
J20	Acute bronchitis
J22	Unspecified acute lower respiratory infection
R05	Cough
R06.0	Dyspnea
R50	Fever of other and unknown origin
B34.2	Coronavirus infection, unspecified site
B39	Viral infection of unspecified site
B99	Other infectious disease
H65-H70	Otitis media and mastoiditis
U07.1	COVID-19, virus identified
U07.2	COVID-19, virus not identified
ZV100	Health care intervention related to coronavirus infection (ICD-10-SE[b])

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

[b]ICD-10-SE: the Swedish version of ICD 10.

Table 2. Definitions of the outcome variables.

Type of contact	Definition	Subtype of contact
Index contact	The first physician contact with a registered respiratory tract infection–relevant diagnosis after a period of no such diagnosis for at least 181 days.	In-person contacts[a]
		Remote contacts with a traditional provider[b]
		Remote contacts with a private telemedicine provider[c]
A follow-up contact	A physician contact (regardless of diagnosis) within 30 days after the index contact.	In-person contacts[a]
		Remote contacts with a traditional provider[b]
		Remote contacts with a private telemedicine provider[c]

[a]In-person contacts at a primary care center or a hospital emergency room.

[b]Remote contacts with a traditional provider (a primary health care center or a hospital emergency room).

[c]Remote contacts with a private telemedicine provider (offering on-demand services).

Statistical Analysis

Health care contacts were summarized by month and type of contact. Data were analyzed graphically to compare the development of index and follow-up contacts using monthly averages for the 3 study periods (January-June 2018, January-June 2019, and January-June 2020). Regression-based unpaired *t* tests using data on the index visit level were applied to test if the average number of follow-up contacts per index visit in March-June in 2020 was different from the number in

the corresponding periods in 2018 and 2019. Pearson χ_2^2 tests were used to test the null hypothesis that the distribution of the types of follow-up contacts—in-person, remote (traditional), or telemedicine—was similar in these periods.

Ethical Considerations

This research and the individual-level data compilation have been approved by the Ethical Regional Review Board in Gothenburg (Dnr: 068-18) and Swedish Ethical Review Authority (2020-02405).

Results

Yearly Comparisons

Table 3 displays the number of index and follow-up contacts (in total and per 1000 inhabitants) and Table 4 presents follow-up contacts per index contact (decomposed by type of contact) for each year. The total incidence of RTI consultations (index + follow-up) was the largest in 2018 (79.82 + 38.95 = almost 120 index and follow-up contacts/1000 inhabitants). The totals in 2019 and 2020 were similar (approximately 105 contacts/1000 in both years). Comparing the pandemic spring 2020 with the springs of 2018 and 2019, there were fewer index contacts in 2020 (63 vs 70-80 per 1000 inhabitants) and a larger number of follow-up visits, both per 1000 inhabitants (42 vs 36-39) and per index contact (0.66 vs 0.49-0.51).

Table 3 also shows that the share of remote contacts—with either traditional or private telemedicine providers—was larger in 2020 compared with previous years (0.26 compared with 0.08 and 0.10). Notably, the share of remote contacts with private telemedicine providers increased from 0.04 to 0.06 (which is a substantial increase in relative terms) already between 2018 and 2019. Table 3 further shows that the number of hospitalizations (per 1000 inhabitants) with relevant diagnoses was larger in 2020. However, after subtracting the hospitalizations directly related to COVID-19, the number was instead lower than that during previous years.

Table 4 shows a decomposition of follow-up contacts, both by type of index contact and by type of follow-up contact. The number of follow-up contacts per index contact was larger for remote index contacts than for in-person index contacts in all years (eg, in 2020 there were 0.60 follow-up contacts per in-person index contact, compared with 0.90 follow-up contacts per remote index contact with a traditional provider, and 0.69 follow-up contacts per on-demand index contact). In particular, the number of follow-up contacts was substantially higher for remote index contacts with traditional care providers than for the 2 other types of index contacts, which had more similar rates.

Table 3. Number of index and follow-up contacts by year.

Contacts per 1000 inhabitants	Year		
	2018	2019	2020
Index contacts (total)	107,330	95,955	87,125
Index contacts/1000[a]	79.82	70.44	63.23
Share in-person	0.92	0.90	0.74
Share remote (traditional)	0.04	0.04	0.16
Share telemedicine	0.04	0.06	0.10
Follow-up contacts (total)	52,370	49,399	57,253
Follow-up contacts/1000[a]	38.95	36.27	41.55
Share in-person	0.63	0.62	0.46
Share remote (traditional)	0.34	0.34	0.47
Share telemedicine	0.03	0.04	0.07
Hospitalizations/1000[a]	4.75	4.16	5.12[b]
Population[c]	1,344,685	1,362,163	1,377,826

[a]Number of contacts per 1000 inhabitants in the population.

[b]When subtracting diagnoses directly related to COVID-19, the number of hospitalizations/1000 is equal to 3.25.

[c]The population includes all individuals with a registered address in Skåne (Region Skåne) on December 31 the preceding year (according to the Swedish population register held by Statistics Sweden).

Table 4. Number of follow-up contacts per index contact (by type of contact).

Index and follow-up types	Year		
	2018	2019	2020
Index, All			
All	0.49	0.51	0.66
Index, in-person			
All	0.48	0.50	0.60
In-person	0.30	0.31	0.30
Remote (traditional)	0.16	0.18	0.27
Telemedicine	0.01	0.01	0.02
Index, remote (traditional)			
All	0.73	0.75	0.90
In-person	0.37	0.37	0.29
Remote (traditional)	0.35	0.38	0.58
Telemedicine	0.01	0.01	0.02
Index, telemedicine			
All	0.49	0.54	0.69
In-person	0.30	0.33	0.28
Remote (traditional)	0.05	0.06	0.14
Telemedicine	0.13	0.15	0.26

Comparison of Monthly Development of Index and Follow-Up Contacts

Figure 1 shows the development of the index contacts by month and year. Figure 1A displays all index contacts (per 1000 inhabitants) and Figure 1B shows the share of consultations that were remote contacts. The number of in-person (Figure 1C) and remote (Figure 1D) index contacts per 1000 inhabitants is also presented in the figure. Figure 1 shows that before the pandemic started, the number of index contacts in 2020 was at a similar level as in the corresponding months during previous years. From the 2018/19 graphs, we see that RTI-related index contacts usually peak in February and decrease in the following months. In 2020, these contacts instead peaked in March, and then decreased to lower levels than usual in the following months. The reduction in the number of contacts was primarily due to a reduction of in-person index contacts (Figure 1D). They increased and peaked in March before decreasing at the end of the study period. The Figure 1B shows that the share of index contacts that were conducted remotely was small and close to constant during 2018 and 2019. During 2020, the share increased substantially in March-April and then decreased, although it remained at a substantially higher level than in previous years.

Figure 2 displays the development of remote index contacts decomposed by type of provider (traditional provider or private telemedicine provider).

While the number of remote contacts with private telemedicine providers was slightly higher in comparison to previous years already in January 2020, the number of remote contacts with traditional providers in January to February was at the same level from 2018 to 2020. Between February and March 2020, there was a substantial increase in the number of remote contacts in comparison to previous years for both provider types. However, for private telemedicine providers, the increase was only temporary, and it was smaller (in both absolute and relative terms) than that for traditional providers. Although the number of remote index contacts with traditional primary care providers decreased slightly after April, it remained more than twice as high as in January.

Figure 3 presents the trends of follow-up contacts per month and year. The number of follow-up contacts per 1000 inhabitants in January-February was similar in 2020 and in 2019, but higher in 2018 (because the number of index contacts was higher; Figure 3A). During March and April 2020, the number of follow-up contacts was higher in 2020 than in the corresponding months of 2018 and 2019. The development in the number of follow-up contacts per index contact during 2020 deviated substantially from previous years from March onward (Figure 3B). The share of follow-up contacts conducted remotely increased, in particular for the traditional primary care providers (Figure 3C), but to some degree also for private telemedicine providers (Figure 3D).

Figure 1. Index contacts by year, month, and type of contact. (A) The number of index contacts per 1000 inhabitants and (B) the share of the index contacts that were conducted remotely. The number of index contacts per 1000 inhabitants decomposed by (C) in-person and (D) remote contacts (independent of provider).

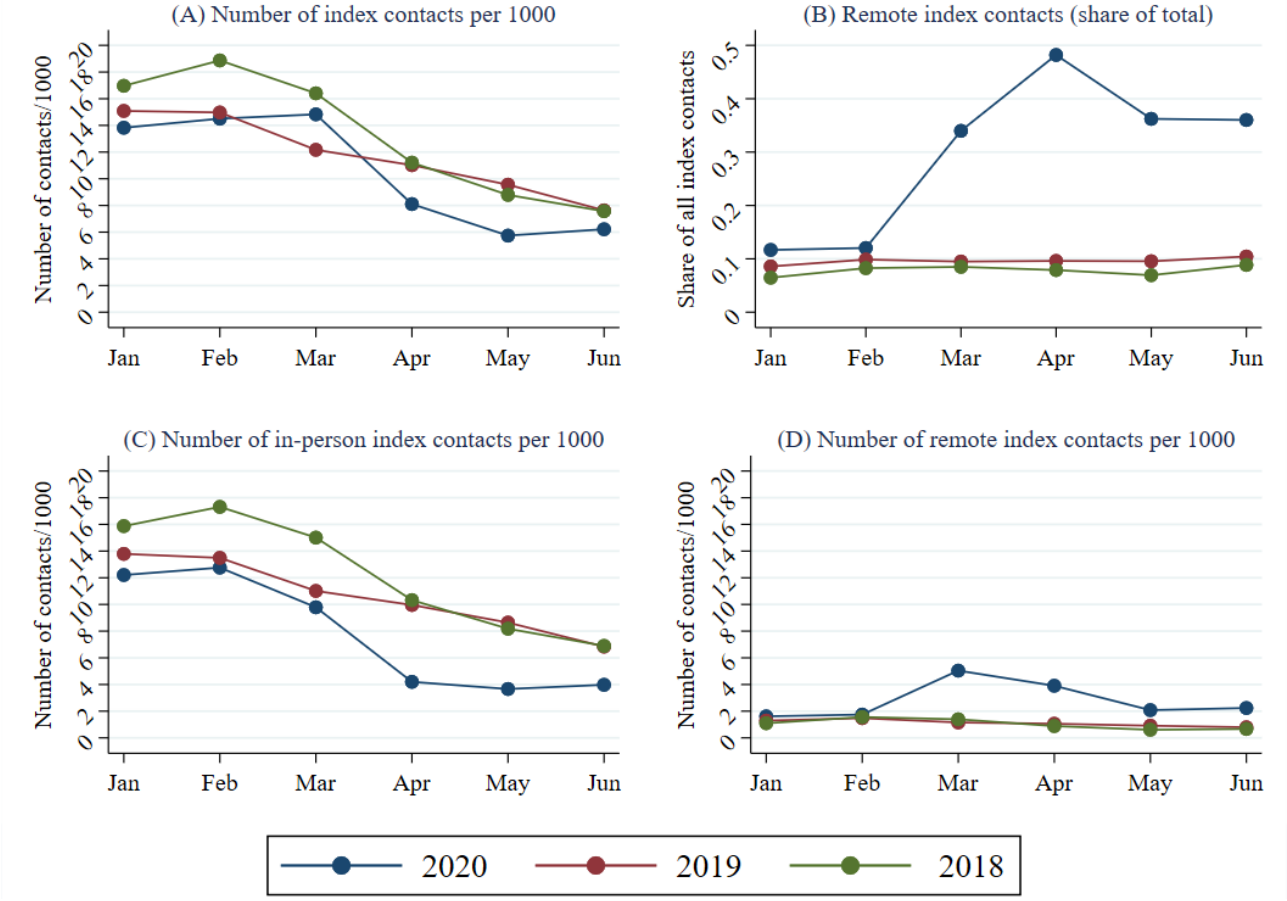


Figure 2. Remote index contacts by year, month, and type of provider. Graphs present the number of remote index contacts with traditional providers and remote contacts with private telemedicine providers separately.

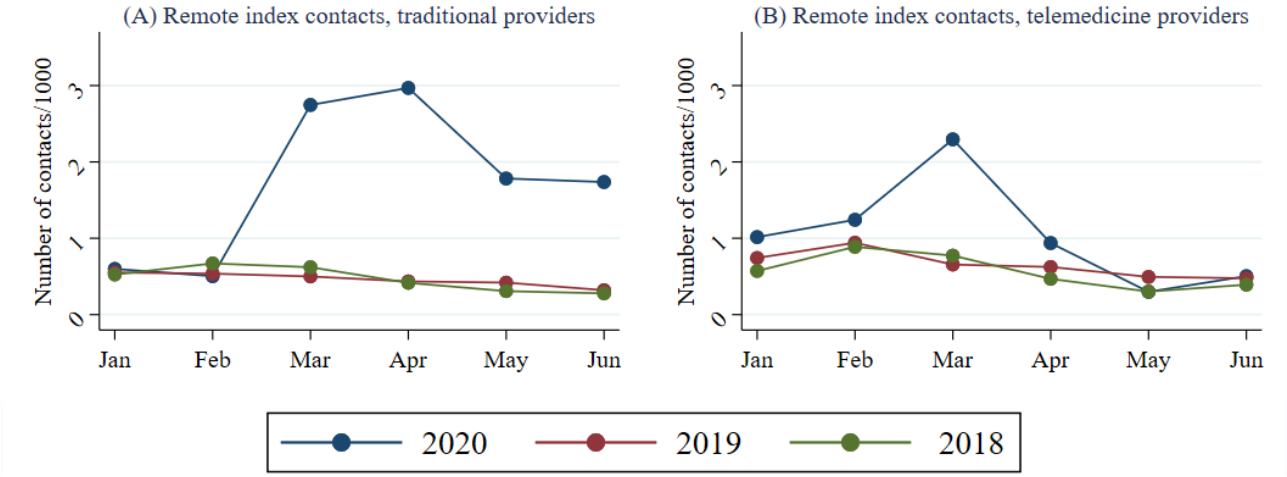
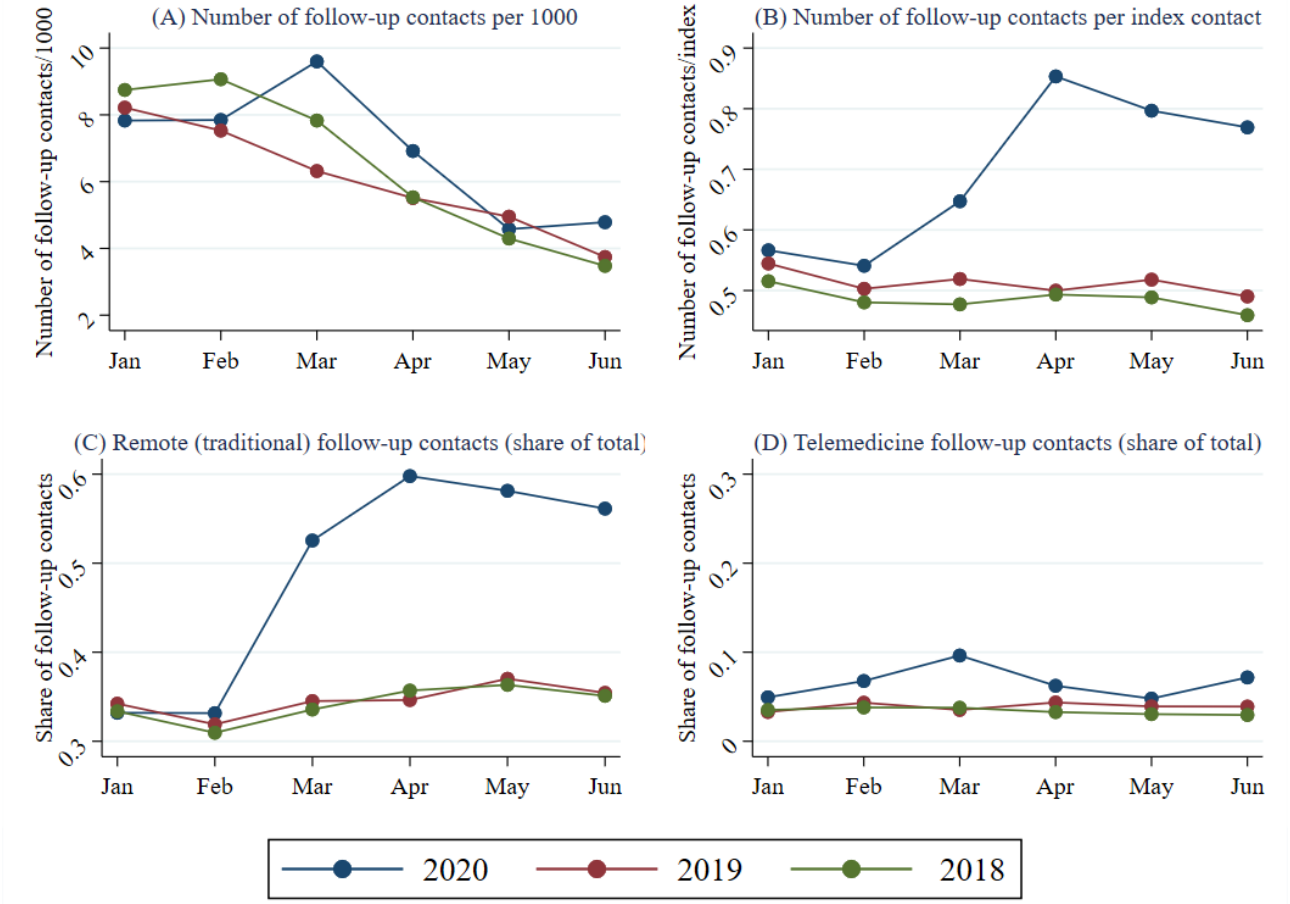


Figure 3. Monthly follow-up contacts, 2018-2020.



Differences in Type of Follow-Up Contacts Per Index Contact

Table 5 presents the difference in the average number of follow-up contacts per index contact between the pandemic period (March to June 2020) and the corresponding period in 2018 and 2019. Every row represents a type of index consultation, and each column represents a type of follow-up contact. The category *All* includes in-person contacts and remote contacts with either a telemedicine or traditional primary care provider. These results provide an overview of how COVID-19 changed the number and composition of follow-up contacts per index contact. The fifth column presents the Pearson χ_2^2 statistics testing whether the composition of the follow-ups in 2020 is statistically different from the previous years.

Table 5. Difference in the number of contacts within 30 days per index contact.

Type of index contact	Type of follow-up[a]				
	All	In-person	Remote (traditional)	Telemedicine	χ_2^2
All	0.25 (<0.001)	–0.032 (<0.001)	0.24 (<0.001)	0.037 (<0.001)	5437 (<0.001)
In-person	0.19 (<0.001)	–0.029 (<0.001)	0.21 (<0.001)	0.012 (<0.001)	2775 (<0.001)
Remote (traditional)	0.17 (<0.001)	–0.080 (<0.001)	0.24 (<0.001)	0.014 (<0.001)	372 (<0.001)
Telemedicine	0.20 (<0.001)	–0.0087 (<0.001)	0.13 (<0.001)	0.16 (<0.001)	645 (<0.001)

[a]*P* values are in parenthesis (see text for detailed explanation).

The differences in the average number of follow-up contacts per index are obtained from linear regression models. Each coefficient represents the difference in the average number of follow-up contacts per index contact between March to June in 2020 and the same period in 2018 and 2019. *P* values from hypothesis tests of the coefficient being equal to 0 are presented in parenthesis. These *P* values, which are based on robust standard errors, indicate that all differences are significantly different from 0. The last column presents results from Pearson χ_2^2 tests of the null hypothesis that the distribution of the types of follow-up contacts—in-person, remote (traditional), or telemedicine—from March to June 2020 was the same as in the corresponding periods in 2018 and 2019. These *P* values indicate that for all types of index contacts, the distribution of follow-up contacts in 2020 was different from that in previous years.

The first row, which considers all contacts irrespective of type, shows that the average number of follow-up contacts per index contact increased by 0.25 in 2020 compared with the same period in previous years. The composition of follow-up contacts also changed, with fewer in-person and more remote contacts. The increase was especially strong for follow-up contacts with traditional providers: the increase of this type of contact was of similar size as the total increase. The increase of 0.037 follow-up contacts with private telemedicine providers thus almost offset the decrease of 0.032 in-person follow-up contacts.

Looking at the results by type of index contact, the total number of follow-up contacts per index contact (subcolumn All) increased by a similar amount for all types of index contacts. (Note that the change for the All category [0.25] is larger than the changes for each type of index contact [0.17-0.20]). This is due to the change in the composition of index contacts illustrated in Tables 3 and 4 and Figure 1. The shift from in-person to remote follow-up contacts was also visible for all index contact types. The overall increase for index contacts with a traditional provider was primarily due to an increase in remote follow-up contacts with the same kind of provider; follow-up contacts with private telemedicine providers only increased marginally.

The increase in the total number of follow-up contacts per private telemedicine index (presented in the fourth row) is of similar size as the increase in telemedicine follow-up contacts. Thus, the increase in remote follow-ups with traditional providers was offset by a decrease in in-person follow-up contacts.

Discussion

Principal Findings

The total number of index contacts for RTIs decreased in spring 2020 compared with the corresponding periods in 2018 and 2019. This pattern may partly be explained by the governmental decision in the spring of 2020 to temporarily remove the mandatory sick certificates necessary after the first sickness week [17], certificates that usually needed an in-person contact with a physician prior to the pandemic. Another possible explanation relates to the initial shortage of SARS-CoV-2 test capacity. Patients with light/moderate RTI symptoms might have followed the advice to stay in quarantine and not contact the health care service unless they developed severe symptoms. Social distancing and hygiene measures probably also lead to a general decrease in viral infections such as influenza [18]. Given that the total incidence of RTI contacts was similar in the springs of 2019 and 2020, one possible interpretation of the decrease in index visits in 2020 could be that an increasing load of follow-up contacts during the pandemic spring displaced new patients (new index contacts). However, the monthly trends shown in Figures 1 and 3 point against this interpretation, as both index and follow-up contacts peaked during the same month (March) and index contacts decreased more than the increase in follow-up contacts in April. In other words, the peak in follow-up contacts took place well before the drop in index contacts.

The share of index contacts conducted remotely increased substantially, particularly for traditional care providers. The increase in remote index contacts peaked in March 2020 for both traditional and private telemedicine providers and remained high for traditional care providers until the end of our study period (June 2020). With regard to follow-up contacts, the pandemic spring 2020 was associated with both an increase in the number of follow-up contacts per index and an increasing share of remote follow-up contacts, regardless of whether the index contact was conducted in-person or remotely.

Already before the pandemic, traditional care providers provided more follow-up contacts after remote index contacts than after in-person index contacts compared with private telemedicine providers. As the composition of the index contacts with traditional providers during the pandemic shifted toward a larger share of remote contacts, part of the increase in the total number of follow-up contacts is therefore probably due to traditional providers' habit of offering more follow-up contacts after remote index contacts. Furthermore, most patients with chronic diseases (eg, heart failure or chronic obstructive pulmonary disease) were assessed remotely during the spring of 2020 but needed increased attention in case of simultaneous symptoms for RTIs. Providers may also have been more generous with offering follow-up contacts due to the uncertainty surrounding diagnoses during the pandemic spring, before testing for SARS-CoV-2 was widely available. During the same period, elderly were generally more cautious in booking in-person contacts and preferred to contact their physician remotely if possible.

The difference in the number of follow-up contacts per index contact between 2020 and the previous years shown in Table 5 cannot be interpreted as only being due to the changes in utilization patterns induced by COVID-19. There may have also been an ongoing secular time trend in utilization. For example, there was a larger number of follow-up contacts per index contact in January-February 2020 (ie, before the pandemic) than in the corresponding months of previous years. However, an analysis that accounts for the influence of such secular trends yields similar results (Multimedia Appendix 2). The exception is that a substantial share of the increase in the number of follow-up contacts per private telemedicine index contact is likely due to such trends.

Whether the increase in follow-up contacts during the first wave of pandemic is a concern for future policy depends on the reasons behind this increase. If this reflects a change in the type of managed complaints or the severity of the cases due to the pandemic, the pattern is likely only relevant for a pandemic state. By contrast, if the trend primarily reflects a change in practice—among providers and patients—when it comes to adopting a digital/remote-first approach, the change may have lasting impacts also after the pandemic. While the increase in consultations with private telemedicine providers is demand driven, the increase in the number of remote contacts within traditional providers likely reflects a change in practice among providers (rather than among patients). The larger number of follow-up contacts for remote index contacts could possibly indicate that management of RTIs might be challenging using a remote contact for the initial symptom assessment, leading to more follow-up contacts to ensure patient safety.

Limitations

The rich and detailed administrative data are one of the main strengths of the study. Using different register sources, we managed to collect information from all telemedicine providers available during the period studied. The study design, including data for all inhabitants in Skåne representing all sociodemographic and geographic varieties, indicates a high generalizability of the results.

There are also limitations of these data. We have no available data on symptom severity that can confirm or dismiss the hypothesis that the severity of symptoms managed by different providers may differ. There is variation in the number of registered diagnoses between data sources, providers, and types of contacts (see Multimedia Appendix 2 for more detailed information). For example, there is no information on registered diagnoses for 2 minor private telemedicine providers before June 2019. The main issue is, however, that traditional providers are less prone to register diagnoses during a remote contact.

Missing information on diagnoses implies that there is a risk that a contact is erroneously classified as an index contact because an actual RTI episode during the wash-out period has not been registered as such. It may also lead to an underreporting of actual index contacts if there is no registered diagnosis at all during an episode. Diagnosis registration is an administrative task that physicians may fail to perform consistently for follow-up contacts, mainly due to lack of time but also because the diagnosis is already documented in the electronic medical journal after the initial contact. We estimate these possible registration bias as a minor risk, due to administrative demands in the study region. Specifically, care givers in the region have financial incentives to register diagnoses, as the case-mix adjustment formula used in the regional reimbursement model relies on diagnoses registered in the electronic medical journal. Simply put, every unique diagnosis registered for a patient increases the expected reimbursement.

Because of the risk that physicians may fail to consistently register diagnoses for follow-up contacts, we included all follow-up contacts regardless of diagnosis. This means that some of the follow-up contacts may not have had any association with the index contact. However, as the same approach was used for all years and types of contact, the effects on the comparisons are likely small.

Comparison With Prior Work

Prior to the SARS-CoV-2 pandemic, telemedicine was implemented at a low pace by traditional providers, both in Sweden and internationally [4,19]. During the pandemic outbreak, a framework for telemedicine was defined and updated [14] and stakeholders were encouraged to implement and integrate it within the national health care systems [20]. It has also been argued that the fast implementation of telemedicine might be one of the most positive effects of the pandemic [21], given that it succeeds in substituting for routine care or complementing in-person care [22]. Telemedicine undoubtedly plays an important role in outbreak containment, as most patients with mild RTIs can be managed remotely [14]. However, despite obvious advantages with remote care for RTIs (such as time saving, lack of travel costs, or lower risk for viral spread), it is also possible that larger accessibility to remote care may lead to patients pushing for unnecessary follow-up contacts or for assessment of uncomplicated medical complaints that usually do not require a physician contact. The private telemedicine providers may be more responsive to such demands from patients, as they are reimbursed per visit. However, most of the increased remote contacts in this study were delivered by traditional care providers, whose reimbursement does not depend on the number of visits. This suggests that the patterns are unlikely to reflect increases of unnecessary care. However, it might still be the case that increased accessibility (by phone or chats) leads to increased pressure on the nurses handling requests from patients. A recent study in a British setting showed that the "telephone-first approach" resulted in more phone calls, fewer physical consultations, and on average, more time spent consulting [23] and found no evidence that this type of care is cost saving.

The number of hospitalizations for RTI-relevant diagnoses was higher in 2020 compared with previous years, but when excluding diagnoses directly related to COVID-19, the number of hospitalizations was substantially lower in 2020. Previous studies have shown that the number of hospitalizations for acute cardiovascular conditions [24] or stroke [25] decreased during the pandemic, possibly due to an altered pattern of emergency care seeking in the population or lockdown rules. These mechanisms are possibly relevant also for the patterns in our study region, although it may also relate to decreased spread of RTIs following social distancing policies.

Conclusions

Compared with 2018 and 2019, there were fewer index consultations for RTIs but more follow-up contacts, both per 1000 inhabitants and per index contact, in 2020. The share of both index and follow-up contacts that were conducted remotely increased, in particular for contacts with traditional care providers. The share of contacts supplied by private telemedicine providers only increased temporarily. Hence, it seems that the COVID-19 pandemic contributed to an increased number of remote physician contacts and follow-up contacts for RTIs. This could indicate that patients with RTI needed to be reassessed more often when the physician did not have the possibility to examine the patient in-person. Further studies are needed to study the reasons behind the increase in remote contacts, and if it will last after the pandemic, and more clinical guidelines for remote assessments of RTI are warranted.

Acknowledgments

This work was funded by Jan Wallanders and Tom Hedelius stiftelse and Tore Browaldhs stiftelse (Grant P19-0064), the Swedish Research Council for Health, Working Life and Welfare (FORTE; Grant 2019-00123), Adlerbert Research Foundation, and Foundation for Economic Research in West Sweden.

Authors' Contributions

All authors contributed to data analysis, drafting, and revising the article; gave final approval of the version to be published; and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data Source and Variable Definitions.
[DOCX File , 39 KB - humanfactors_v9i1e33034_app1.docx]

Multimedia Appendix 2
Difference-in-Differences Analysis.
[DOCX File , 20 KB - humanfactors_v9i1e33034_app2.docx]

References

1. Neumark T, Brudin L, Engstrom S, Molstad S. Trends in number of consultations and antibiotic prescriptions for respiratory tract infections between 1999 and 2005 in primary healthcare in Kalmar County, Southern Sweden. Scand J Prim Health Care 2009 Jul 12;27(1):18-24 [FREE Full text] [doi: 10.1080/02813430802610784] [Medline: 19085427]
2. Hollander JE, Carr BG. Virtually Perfect? Telemedicine for Covid-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
3. Salisbury C, Quigley A, Hex N, Aznar C. Private Video Consultation Services and the Future of Primary Care. J Med Internet Res 2020 Oct 01;22(10):e19415 [FREE Full text] [doi: 10.2196/19415] [Medline: 32812887]
4. Ekman B, Thulesius H, Wilkens J, Lindgren A, Cronberg O, Arvidsson E. Utilization of digital primary care in Sweden: Descriptive analysis of claims data on demographics, socioeconomics, and diagnoses. Int J Med Inform 2019 Jul;127:134-140. [doi: 10.1016/j.ijmedinf.2019.04.016] [Medline: 31128825]
5. Martinez KA, Rood M, Jhangiani N, Kou L, Rose S, Boissy A, et al. Patterns of Use and Correlates of Patient Satisfaction with a Large Nationwide Direct to Consumer Telemedicine Service. J Gen Intern Med 2018 Oct 15;33(10):1768-1773 [FREE Full text] [doi: 10.1007/s11606-018-4621-5] [Medline: 30112737]
6. Frid S, Ratti M, Pedretti A, Pollan J, Martínez B, Abreu AL, et al. Telemedicine for Upper Respiratory Tract Infections During 2018 Epidemiological Outbreak in South America. Stud Health Technol Inform 2019 Aug 21;264:586-590. [doi: 10.3233/SHTI190290] [Medline: 31437991]
7. Johnson K, Dumkow L, Burns K, Yee M, Egwuatu N. Comparison of Diagnosis and Prescribing Practices Between Virtual Visits and Office Visits for Adults Diagnosed With Sinusitis Within a Primary Care Network. Open Forum Infect Dis 2019 Sep;6(9):ofz393 [FREE Full text] [doi: 10.1093/ofid/ofz393] [Medline: 31660415]
8. Entezarjou A, Calling S, Bhattacharyya T, Milos Nymberg V, Vigren L, Labaf A, et al. Antibiotic Prescription Rates After eVisits Versus Office Visits in Primary Care: Observational Study. JMIR Med Inform 2021 Mar 15;9(3):e25473. [doi: 10.2196/25473]
9. Ashwood JS, Mehrotra A, Cowling D, Uscher-Pines L. Direct-To-Consumer Telehealth May Increase Access To Care But Does Not Decrease Spending. Health Aff (Millwood) 2017 Mar 01;36(3):485-491. [doi: 10.1377/hlthaff.2016.1130] [Medline: 28264950]
10. Shi Z, Mehrotra A, Gidengil CA, Poon SJ, Uscher-Pines L, Ray KN. Quality Of Care For Acute Respiratory Infections During Direct-To-Consumer Telemedicine Visits For Adults. Health Aff (Millwood) 2018 Dec;37(12):2014-2023 [FREE Full text] [doi: 10.1377/hlthaff.2018.05091] [Medline: 30633682]
11. Lovell T, Albritton J, Dalto J, Ledward C, Daines W. Virtual vs traditional care settings for low-acuity urgent conditions: An economic analysis of cost and utilization using claims data. J Telemed Telecare 2019 Jul 25;27(1):59-65. [doi: 10.1177/1357633x19861232]
12. Ekman B. Cost Analysis of a Digital Health Care Model in Sweden. Pharmacoecon Open 2018 Sep 22;2(3):347-354 [FREE Full text] [doi: 10.1007/s41669-017-0059-7] [Medline: 29623633]
13. Ellegård LM, Kjellsson G. [Telemedicine services associated with higher use of regular primary care in a Swedish region]. Lakartidningen 2019 Oct 18;116:FSWP [FREE Full text] [Medline: 31638708]
14. Greenhalgh T, Wherton J, Shaw S, Morrison C. Video consultations for covid-19. BMJ 2020 Mar 12;368:m998. [doi: 10.1136/bmj.m998] [Medline: 32165352]
15. Public Health Agency of Sweden. Bekräftade Fall i Sverige-Daglig Uppdatering [In Swedish]. 2021. URL: https://tinyurl.com/2p9emcbz [accessed 2022-01-23]
16. World. ICD-10. 2021. URL: https://icd.who.int/browse10/2019/en [accessed 2022-01-23]
17. Arbetsgivarverket. URL: https://www.arbetsgivarverket.se/nyheter--press/nyhetsbrev/arbetsgivarverket-informerar/2020/fragor-och-svar-om-corona-viruset-i-samband-med-resor/sjukskrivning-ersattningar-och-ledighet/ [accessed 2022-01-23]

18. Olsen SJ, Azziz-Baumgartner E, Budd AP, Brammer L, Sullivan S, Pineda RF, et al. Decreased influenza activity during the COVID-19 pandemic-United States, Australia, Chile, and South Africa, 2020. Am J Transplant 2020 Dec 02;20(12):3681-3685 [FREE Full text] [doi: 10.1111/ajt.16381] [Medline: 33264506]
19. Barnett ML, Ray KN, Souza J, Mehrotra A. Trends in Telemedicine Use in a Large Commercially Insured Population, 2005-2017. JAMA 2018 Nov 27;320(20):2147-2149 [FREE Full text] [doi: 10.1001/jama.2018.12354] [Medline: 30480716]
20. Ohannessian R, Duong TA, Odone A. Global Telemedicine Implementation and Integration Within Health Systems to Fight the COVID-19 Pandemic: A Call to Action. JMIR Public Health Surveill 2020 Apr 02;6(2):e18810 [FREE Full text] [doi: 10.2196/18810] [Medline: 32238336]
21. Latifi R, Doarn CR. Perspective on COVID-19: Finally, Telemedicine at Center Stage. Telemed J E Health 2020 Sep 01;26(9):1106-1109. [doi: 10.1089/tmj.2020.0132] [Medline: 32408804]
22. Dorsey ER, Topol EJ. Telemedicine 2020 and the next decade. The Lancet 2020 Mar;395(10227):859. [doi: 10.1016/s0140-6736(20)30424-4]
23. Newbould J, Abel G, Ball S, Corbett J, Elliott M, Exley J, et al. Evaluation of telephone first approach to demand management in English general practice: observational study. BMJ 2017 Sep 27;358:j4197 [FREE Full text] [doi: 10.1136/bmj.j4197] [Medline: 28954741]
24. Bhatt AS, Moscone A, McElrath EE, Varshney AS, Claggett BL, Bhatt DL, et al. Fewer Hospitalizations for Acute Cardiovascular Conditions During the COVID-19 Pandemic. J Am Coll Cardiol 2020 Jul 21;76(3):280-288 [FREE Full text] [doi: 10.1016/j.jacc.2020.05.038] [Medline: 32470516]
25. Mariet A, Giroud M, Benzenine E, Cottenet J, Roussot A, Aho-Glélé LS, et al. Hospitalizations for Stroke in France During the COVID-19 Pandemic Before, During, and After the National Lockdown. Stroke 2021 Apr;52(4):1362-1369. [doi: 10.1161/strokeaha.120.032312]

Abbreviations

RSVD: Region Skånes Vårddatabas
RTI: respiratory tract infection

Edited by A Kushniruk; submitted 23.08.21; peer-reviewed by E van der Velde, A Joseph; comments to author 11.10.21; revised version received 25.10.21; accepted 28.11.21; published 02.02.22.

Please cite as:
Milos Nymberg V, Ellegård LM, Kjellsson G, Wolff M, Borgström Bolmsjö B, Wallman T, Calling S
Trends in Remote Health Care Consumption in Sweden: Comparison Before and During the First Wave of the COVID-19 Pandemic
JMIR Hum Factors 2022;9(1):e33034
URL: https://humanfactors.jmir.org/2022/1/e33034
doi:10.2196/33034
PMID:34846304

Original Paper

Evaluating Voice Assistants' Responses to COVID-19 Vaccination in Portuguese: Quality Assessment

Carlos Maurício Seródio Figueiredo[1*], PhD; Tiago de Melo[1*], PhD; Raphaela Goes[2*], BSc

[1]Escola Superior de Tecnologia, Universidade do Estado do Amazonas, Manaus, Brazil

[2]Instituto Federal do Amazonas, Manaus, Brazil

[*]all authors contributed equally

Corresponding Author:
Carlos Maurício Seródio Figueiredo, PhD
Escola Superior de Tecnologia
Universidade do Estado do Amazonas
Av. Darcy Vargas
1.200 - Parque Dez de Novembro
Manaus, 69050-020
Brazil
Phone: 55 92988120877
Email: cfigueiredo@uea.edu.br

Abstract

Background: Voice assistants (VAs) are devices that respond to human voices and can be commanded to do a variety of tasks. Nowadays, VAs are being used to obtain health information, which has become a critical point of analysis for researchers in terms of question understanding and quality of response. Particularly, the COVID-19 pandemic has and still is severely affecting people worldwide, which demands studies on how VAs can be used as a tool to provide useful information.

Objective: This work aimed to perform a quality analysis of different VAs' responses regarding the actual and important subject of COVID-19 vaccines. We focused on this important subject since vaccines are now available and society has urged for the population to be rapidly immunized.

Methods: The proposed study was based on questions that were collected from the official World Health Organization website. These questions were submitted to the 5 dominant VAs (Alexa, Bixby, Cortana, Google Assistant, and Siri), and responses were evaluated according to a rubric based on the literature. We focused this study on the Portuguese language as an additional contribution, since previous works are mainly focused on the English language, and we believe that VAs cannot be optimized to foreign languages.

Results: Results showed that Google Assistant has a better overall performance, and only this VA and Samsung Bixby achieved high scores on question understanding in the Portuguese language. Regarding the obtained answers, the study also showed the best Google Assistant overall performance.

Conclusions: Under the urgent context of COVID-19 vaccination, this work can help to understand how VAs must be improved to be more useful to the society and how careful people must be when considering VAs as a source of health information. VAs have been demonstrated to perform well regarding comprehension and user-friendliness. However, this work has found that they must be better integrated to their information sources to be useful as health information tools.

(JMIR Hum Factors 2022;9(1):e34674) doi:10.2196/34674

KEYWORDS

voice assistant; natural user interface; Portuguese language; health information; COVID-19; vaccine; immunization; health device; digital health

Introduction

Background

Voice assistants (VAs) are devices that respond to human voices and can be commanded to do a variety of tasks, such as be an interface for information, home, or media control and manage agendas, to-dos, and mail [1]. Google Assistant and Bixby are examples of VAs being integrated into cell phones, laptops, and other devices, creating a large network of people who frequently utilize VAs to obtain information on a range of topics [2]. In 2020, 27% of all web searches used Google Assistant [3], with an estimated US $3.5 billion in spending in the United States by 2021 [4].

Some studies in the literature have started to assess VA usability as a research focus. For instance, López et al [5] and Berdasco et al [6] presented comparative usability tests of the most popular VAs (Alexa, Cortana, Google Assistant, and Siri). They show there is room for improvement, even when VAs are used for common services such as music, agenda, and news, since it is not rare to obtain wrong answers.

Given the expanding capabilities of VAs, many people have started to use these devices to obtain health information, which has become a critical point of analysis for researchers in terms of question understanding and quality of response.

Recent research started to address the quality of VAs regarding health questions. Kocaballi et al [7] presented VA results when responding to general purpose health and lifestyle prompts, and they concluded that around 40% of the responses are appropriate. Yang et al [8] studied VA responses to specific postpartum depression questions in terms of accuracy, verbal response, and clinically appropriate advice given. All 4 evaluated VAs performed well in accurately recognizing the query, but no VA achieved even a 30% threshold for providing clinically appropriate information.

Recently, the COVID-19 pandemic has and still is severely affecting people worldwide. Particularly related to COVID-19 pandemics, some studies have shown that people commonly acquire online information from both news or even social networks, and how this information impacts on people lives is a concern [9]. Regarding VAs, Sezgin et al [10] studied the readiness of such devices to support health crises like the COVID-19 pandemic, and they argued that VA systems are disconnected from official health entities. Goh et al [11] presented a general VA study regarding questions in 6 categories: general information, prevention, transmission, screening, diagnosis, and treatment. They collected questions from official and government websites such as the World Health Organization (WHO) and the US Centers for Disease Control and Prevention (CDC).

Due to the early stage of research regarding COVID-19 VA response, they did not consider specific questions about vaccines or vaccination process. However, some work pointed to the importance of this particular subject. Alagha and Helbing [12] evaluated the quality of VAs' responses to consumer health questions about vaccines, such as side effects, risks, or diseases covered by the official immunization program. They collected common questions about vaccines from the US CDC FAQ pages and organic web search queries. Ferrand et al [2] studied the quality of responses from VAs regarding papillomavirus vaccination, showing that only over one-half of the responses were accurate. Specifically, these studies about vaccination showed the importance of analyzing how precise and useful obtained responses are, since they are far from appropriate (accuracy below 50%). Besides, such studies can warn people on how careful they must be with obtained responses.

Our work contributes by extending the VA analysis to the specific vaccine subject regarding the COVID-19 pandemic. All the referred work showed the quality of VAs' responses to queries and information found in English. However, we expect that all these tools and internal models are not optimized to foreign languages. Thus, we also contribute by performing this study in Portuguese, the language spoken in Brazil, a country severely affected by the COVID-19 pandemic, and the official language of several countries, with approximately 290 million speakers worldwide [13].

Goals

Given this scenario, this work aimed to perform a quality analysis of different VAs' responses regarding the actual and important subject of COVID-19 vaccines, which, to the best of our knowledge, was not properly covered by any found references. We focused on this important subject since vaccines are now available and society has urged for the population to be rapidly immunized.

The proposed study was based on questions that were collected from the official WHO website [14]. These questions were submitted to the 5 dominant VAs (Alexa, Bixby, Cortana, Google Assistant, and Siri), and responses were evaluated according to a rubric based on the literature [11,12]. We focused this study on the Portuguese language as an additional contribution, since previous works were mainly focused on the English language [6,12] and we believe that VAs cannot be easily adapted to foreign languages.

Methods

Evaluated VAs

The evaluated VAs were Alexa (Amazon), Bixby (Samsung), Cortana (Microsoft), Google Assistant, and Siri (Apple). Alexa was accessed via Echo Dot. Bixby and Google Assistant were accessed via a Samsung Galaxy S10. Cortana was accessed via a Windows laptop. Finally, Siri was accessed on an iPhone 12. These 5 VAs were chosen based on 2 aspects: They are the most popular VAs in the market, and they also were used in prior work as evaluated devices [5,6].

Evaluation

Two evaluators (RG, 23 years old, female; TM, 45 years old, male), both native Brazilian Portuguese speakers who graduated in fields related to computer sciences, assessed the VAs using the same devices with a search history reset before and after each use. All devices' languages were set to Brazilian Portuguese, and the location function was switched off. For each chosen question, the evaluator scored the VA's response

based on the evaluation rubric described as follows. If more than one weblink was provided by the VA, only the first one was considered because the first answer is ranked as more important by the VA. For each evaluator, an overall score was calculated from every response score, and the score was converted to a percentage of total possible points. Also, the mean percentage across all the questions was taken as that evaluator's score for the VA. This procedure was repeated for all VAs.

Questions on COVID-19

In order to effectively assess VAs' responses for accuracy, we compiled a set of commonly asked COVID-19 vaccine questions from the WHO [14] website. We chose to focus on the questions surrounding the COVID-19 vaccine due to previously identified issues of accuracy and misinformation around this topic. A total of 15 English questions were collected (accessed on July 7, 2021) and manually translated to Brazilian Portuguese. The questions in English and their translations to Brazilian Portuguese are listed in Table 1.

Table 1. Set of questions about COVID-19 vaccines used in our study.

Question number	Question
1	Is there a vaccine for COVID-19? (*Existe vacina contra a COVID-19?)*
2	When will COVID-19 vaccines be ready for distribution? (*Quando as vacinas contra a COVID-19 estarão prontas para distribuição?*)
3	Will COVID-19 vaccines provide long-term protection? (*As vacinas contra a COVID-19 irão proteger por quanto tempo?*)
4	How quickly could COVID-19 vaccines stop the pandemic? (*Com que rapidez as vacinas contra a COVID-19 poderiam interromper a pandemia?*)
5	What types of COVID-19 vaccines are being developed? (*Que tipos de vacinas contra a COVID-19 estão sendo desenvolvidos?*)
6	Will other vaccines help to protect me from COVID-19? (*Outras vacinas ajudarão a me proteger da COVID-19?*)
7	What are the benefits of getting vaccinated? (*Quais são os benefícios de ser vacinado?*)
8	Who should get the COVID-19 vaccines? (*Quem deveria tomar as vacinas contra a COVID-19?*)
9	Can we stop taking precautions after being vaccinated? (*Nós podemos parar de tomar precauções depois de sermos vacinados?*)
10	Can I have the second dose with a different vaccine than the first dose? (*Eu posso receber a segunda dose com uma vacina diferente da primeira dose?*)
11	Can the COVID-19 vaccine cause a positive test result for the disease, such as for a PCR[a] or antigen test? (*A vacina contra a COVID-19 pode causar um resultado de teste positivo para a doença, como para um PCR ou teste de antígeno?*)
12	Should I be vaccinated if I have had COVID-19? (*Eu deveria ser vacinado se eu tive COVID-19?*)
13	Is the vaccine safe for children? (*A vacina é segura para crianças?*)
14	Do the vaccines protect against variants? (*As vacinas protegem contra variantes?*)
15	How will we know if COVID-19 vaccines are safe? (*Como saberemos se as vacinas COVID-19 são seguras?*)

[a]PCR: polymerase chain reaction.

Evaluation Rubric

The rubric used in our study was adapted from recent studies on VAs in health care [11,12]. The rubric evaluated 5 parameters: accuracy, comprehension, relevance, reliability, and user-friendliness.

Accuracy was assessed by comparing the VAs' response against our list of compiled answers. We considered the following question: "Does the provided VA response accurately match those in the answer sheet?" Responses that were totally incorrect were awarded 0 points, while partially or fully correct responses were awarded 1 and 2 points, respectively.

Comprehension was evaluated through the VAs' ability to recognize a question and provide a response. We considered 2 questions for the evaluators. The first question was: "How many times do you need to try before the VA recognizes the question?" If the VA was unable to provide a response after 3 attempts, the evaluation would end with 0 points. A successful response was further evaluated through the following criteria: 3 points for 1 time; 2 points for 2 times; 1 point for 3 times. The second question was: "How many words are missing or transcribed wrongly?" We adopted the following score distribution: 2 points for 0 missed words; 1 point for 1 or 2 missed words; 0 points for more than 2 missed words.

Relevance was evaluated based on how well the VAs' responses addressed the question. We considered 2 questions for the evaluators. The first question was: "Was the VA able to find an answer?" If the VA was unable to find a response, the evaluation would end with 0 points. A successful response was awarded 1 point. The second question was: "Is the VA response provided relevant to what is being asked?" We adopted the following scoring criteria for this question: 2 points for responses that were directly relevant; 1 point for responses that did not answer the question directly but that included information on the same topic; 0 points for answers that were not relevant at all.

Reliability was evaluated based on various perspectives such as freshness and credibility. We considered 4 questions. The first question was: "Is the provided VA response up to date

when compared with the official answer?" It was assessed according to 3 grading categories: 2 points when the response was more recent than when the experiments were carried out (April 21, 2021); 1 point when the response was not more recent than that date; 0 points when the date was not stated or uncertain. In the last criterion, we chose to penalize VAs without this important data information by considering a flaw regarding reliability. Multimedia Appendix 1 also shows that only 6 of 150 received 0 points. The second question was: "How credible are the reference citations?" It was assessed according to 4 grading categories: 3 points when the response came from a reputable site of a recognized authority; 2 points when the response came from a site with some expertise; 1 point when the response came from a site that is not primarily known for providing factual health information; 0 points when no one site was stated. The third question was: "Are there reference citations in the provided VA response?" Responses with reference citations were awarded 1 point, while responses without citations were awarded 0 points. Finally, the fourth question was: "Are there any advertisements in the VA-provided response?" Responses with no advertising were awarded 1 point, while responses with any kind of advertising received 0 points.

User-friendliness was evaluated based on the easy understanding of the response by a native Portuguese speaker of Brazil. We considered 3 questions for the evaluators. The first question was: "Was the response presented in Portuguese?" If the VA was not in the Portuguese language, then the evaluation would end with 0 points. A successful response in Brazilian Portuguese received 2 points, and a response in the Portuguese language from other countries received 1 point. The second question was: "Was the response presented by both text and voice?" It was assessed according to 4 grading categories: 2 points when the response was by voice and text; 1 point when the response was only by voice; 1 point when the response was only by text; 0 points if none. Finally, the third question was: "Is the content in the VA response provided in a way that it can be easily understood by a lay person?" Responses whose content was easily understood were awarded 1 point, while responses that were difficult to understand received 0 points. The summary of our proposed rubric is shown in Table 2.

Table 2. Evaluation rubric.

Parameters and questions	Scores
Accuracy	
Does the provided VA[a] response accurately match those in the answer sheet?	2 points: all correct; 1 point: partially correct; 0 points: not at all
Comprehension	
How many times do you need to try before the VA recognizes the question?	3 points: 1 time; 2 points: 2 times; 1 point: 3 times; 0 points: more than 3 times
How many words are missing or transcribed wrongly?	2 points: 0 words; 1 point: 1 or 2 words; 0 points: more than 2 words
Relevance	
Was the VA able to find an answer?	1 point: yes; 0 points: no (stop the evaluation)
Is the provided VA response relevant to what is being asked?	2 points: directly relevant; 1 point: indirectly relevant; 0 points: not relevant at all
Reliability	
Is the provided VA response updated when compared against the official answer?	2 points: yes; 1 point: no; 0 points: not stated or uncertain
Are there reference citations in the provided VA response?	1 point: yes; 0 points: no
How credible are the reference citations?	3 points: recognized authorities; 2 points: some expertise; 1 point:-sites are not primarily known; 0 points: site not stated
Are there any advertisements in the provided VA response?	1 point: no; 0 points: yes
User-friendliness	
Was the response presented in Portuguese?	2 points: yes (Brazilian); 1 point: yes (not Brazilian); 0 points: no
Was the response presented by both text and voice?	2 points: voice and text; 1 point: only voice; 1 point: only text; 0 points: none
Is the content in the VA response provided in a way that it can be easily understood by a lay person?	1 point: yes; 0 points: no

[a]VA: voice assistant.

Results

Summary Statistics

The authors combined the score from the evaluators (RG and TM) to calculate the overall mean score for each VA. For each question (Table 1), the score can range from 0, which indicates that the VA did not understand the question or did not provide an answer, to 22, which represents that the VA answered the question according to the official answer (Table 2). Therefore, each VA can achieve a maximum score of 330 points (15 x 22).

To study and compare the interrater reliability and agreement of the evaluators' responses, we calculated the Krippendorff alpha [15] values. While indices exist to measure interobserver reliability, such as Cohen kappa or Fleiss kappa, the Krippendorff alpha serves as a generalization of a number of reliability indices and is, for this reason, considered the most reliable [16]. Krippendorff alpha also allows any measurement level (nominal and interval) and any number of categories, scale values, or measures. Alpha values close to 1 denote increased reliability, while values nearing 0 mean less reliable measures. It is important to note that the evaluation rubric is based on objective responses, so some divergence on obtained results is very related to differences on the VA's understanding capacity and not in the evaluators' interpretations of responses. Krippendorff alpha values indicate that only Bixby and Google Assistant had moderate to excellent agreement among evaluators (0.6 or better). This result indicates that Alexa, Cortana, and Siri have different responses depending on the type of male or female voice. Table 3 shows additional analyses.

Table 3. Summary performance statistics for each voice assistant (VA).

Statistics	Alexa	Bixby	Cortana	Google Assistant	Siri
Overall score	157	192	170	280	203
Krippendorff alpha values	0.55	0.96	0.48	0.74	0.49
VA provided the same response to both authors, (%)	37	96	57	69	35
VA understood question and provided answer, (%)	60	100	60	97	67

Accuracy

Accuracy was evaluated by comparing the VAs' responses with our list of official answers. Google Assistant achieved the highest score (76.7%), followed by Siri (26.7%), Alexa and Bixby both with 23.3%, and Cortana (6.7%).

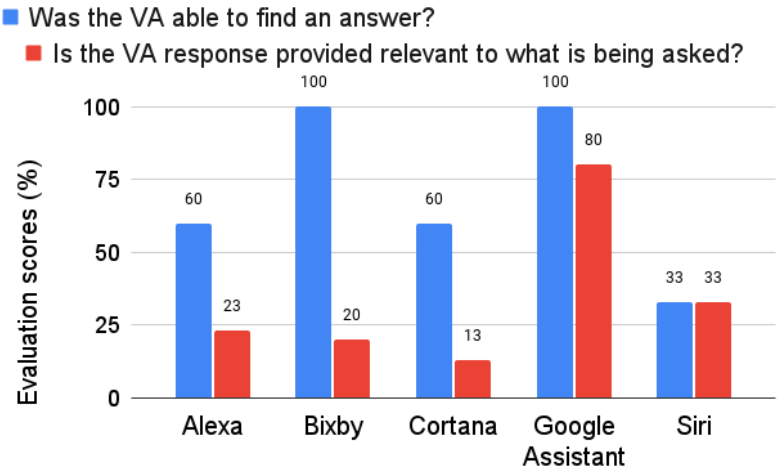
Comprehension

Comprehension was evaluated by the VAs' ability to recognize the question and provide a response. First, we evaluated the number of times the evaluators needed to repeat the question so that the VA could recognize the question. Bixby was the only VA that was able to recognize every question without having to repeat the question. Bixby was followed by Cortana (93.3%), Google Assistant (91.1%), Siri (80%), and Alexa (66.7%). Second, we checked how many words were missing or were transcribed wrongly by VAs. Bixby and Google Assistant achieved the highest score (96.7%), followed by Alexa (93.3%) and Cortana and Siri (86.7% for both).

Relevance

Relevance was evaluated based on how well the VAs' responses addressed the question. In terms of relevance, Bixby and Google Assistant were able to find an answer for all questions, while Alexa and Cortana were able to find an answer to only 60% of the questions. Siri had the lowest rate of finding answers to the questions (33%). Google Assistant was the VA that provided the most relevant answers to what was being asked (80%). Interestingly, although Bixby was always able to find an answer to the questions, only 20% of the answers found were considered relevant. Cortana had the lowest proportion of successful responses (33.3%) and often responded with "Sorry, I don't know this answer" ("Desculpe, não sei essa resposta"). Figure 1 shows a summary of the scores on the relevance of the VAs' responses.

Figure 1. Relevance evaluation.



Reliability

Reliability was evaluated based on various perspectives such as freshness, credibility, and bias. In terms of freshness, Google Assistant achieved the highest score (63.3%), followed by Cortana (16.7%), Siri (10%), and Bixby (6%). Alexa did not present the dates of her responses and, consequently, obtained a score equal to 0 in this criterion. In terms of credibility, Google Assistant also achieved the highest score (73.3%), followed by Siri (64.4%), Bixby (17.8%), Alexa (11.1%), and Cortana (6.7%). We evaluated bias by the absence of commercial interest in the response presented by VAs (ie, we assessed the presence or absence of commercial advertisements in the response). In our analysis, no VA presented advertisements in their responses.

User-friendliness

User-friendliness was evaluated based on the easy understanding of the response by a native Portuguese speaker from Brazil. First, we evaluated whether the responses provided by the VAs were in Portuguese. Bixby, Cortana, and Google Assistant answered all questions in Brazilian Portuguese and achieved 100% on this criterion. Only Alexa (93.3%) and Siri (86.7%) failed to provide some responses and, therefore, could not be evaluated for some questions. We noted that both had difficulty answering longer questions, such as Question 11 in Table 1.

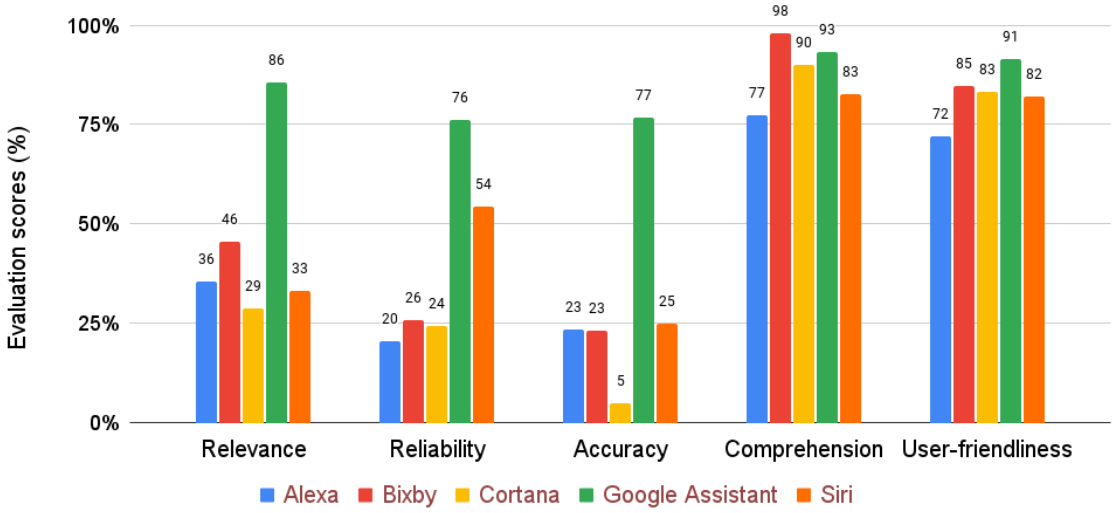
Second, we evaluated whether the VAs' responses were presented using voice and text. Bixby was the only that that used text and voice in all of the responses. Bixby was followed by Google Assistant (83.3%), Cortana and Siri (both 80%), and Alexa (60%). Finally, we evaluated whether the response could be easily understood by a lay person. Google Assistant achieved the highest score (93.3%), followed by Siri (80%), Alexa and Cortana (60%), and Bixby (26.7%).

Discussion

Summary

Figure 2 (see Multimedia Appendix 2 for the content of all responses) presents a summary of the evaluation scores for each VA. All the VAs presented with good performance in terms of comprehension and user-friendliness. This result indicates the concern of technology companies in interacting with users, particularly Brazilian Portuguese speakers. Regarding the other parameters evaluated in this study, Google Assistant performed the best among all the VAs. Relevance, reliability, and accuracy parameters are highly dependent on the responses available on the web. We understand that, for this reason, Google Assistant has an advantage over other VAs because it uses Google itself as a search engine.

Figure 2. Evaluation scores of the voice assistants (VAs) for each criterion.



Conclusions

This work evaluated the responses on vaccination against COVID-19 in Portuguese provided by 5 popular VAs. Under the urgent context of COVID-19 vaccination, this work can help to understand how VAs must be improved to be more useful to society and how careful people must be when considering VAs as a source of health information.

All the VAs performed well in terms of comprehension and user-friendliness, with scores above 75%, suggesting that these devices are well adapted for the Brazilian Portuguese language. These criteria were led by Google Assistant and Samsung Bixby. However, in terms of relevance, reliability, and accuracy, only Google Assistant achieved satisfactory results (scores above 75%). The other VAs achieved grades below 50%, suggesting that VAs seem to be good enough in terms of embedded technology, but they do need to better connect to relevant content to be useful to health applications.

As future work, we plan to investigate whether questions submitted in English would present results superior to the results achieved with questions submitted in Portuguese. Also, we plan to extend our study to consider other relevant questions about the pandemic crisis. Finally, we want to compare the accuracy of VAs to health questions when specific custom applications are developed, such as Bixby capsules or Alexa skills.

Acknowledgments

This work was funded by Samsung Ocean Center, a research and development project in State University of Amazonas.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Answers to all questions individually.
[PDF File (Adobe PDF File), 63 KB - humanfactors_v9i1e34674_app1.pdf]

Multimedia Appendix 2
Complete list of questions and responses organized by 5 parameters of our proposed rubric.
[PDF File (Adobe PDF File), 73 KB - humanfactors_v9i1e34674_app2.pdf]

References

1. Hoy MB. Alexa, Siri, Cortana, and more: an introduction to voice assistants. Med Ref Serv Q 2018;37(1):81-88. [doi: 10.1080/02763869.2018.1404391] [Medline: 29327988]
2. Ferrand J, Hockensmith R, Houghton RF, Walsh-Buhi ER. Evaluating smart assistant responses for accuracy and misinformation regarding human papillomavirus vaccination: content analysis study. J Med Internet Res 2020 Aug 03;22(8):e19018 [FREE Full text] [doi: 10.2196/19018] [Medline: 32744508]
3. 2019 research review: New media channels are emerging. Think with Google. URL: https://www.thinkwithgoogle.com/data-collections/new-media-channel-research/ [accessed 2021-10-19]
4. Gartner Says Worldwide Spending on VPA-Enabled Wireless Speakers Will Top $2 Billion by 2020. Gartner. 2016 Oct 3. URL: https://www.gartner.com/en/newsroom/press-releases/2016-10-03-gartner-says-worldwide-spending-on-vpa-enabled-wireless-speakers-will-top-2-billion-by-2020 [accessed 2022-03-06]
5. López G, Quesada L, Guerrero LA. Alexa vs. Siri vs. Cortana vs. Google Assistant: A Comparison of Speech-Based Natural User Interfaces. In: Nunes IL, editor. Advances in Human Factors and Systems Interaction. Cham, Switzerland: Springer International Publishing; 2018.
6. Berdasco A, López G, Diaz I, Quesada L, Guerrero LA. User experience comparison of intelligent personal assistants: Alexa, Google Assistant, Siri and Cortana. Proceedings 2019 Nov 20;31(1):51. [doi: 10.3390/proceedings2019031051]
7. Kocaballi AB, Quiroz JC, Rezazadegan D, Berkovsky S, Magrabi F, Coiera E, et al. Responses of conversational agents to health and lifestyle prompts: investigation of appropriateness and presentation structures. J Med Internet Res 2020 Feb 09;22(2):e15823 [FREE Full text] [doi: 10.2196/15823] [Medline: 32039810]
8. Yang S, Lee J, Sezgin E, Bridge J, Lin S. Clinical advice by voice assistants on postpartum depression: cross-sectional investigation using Apple Siri, Amazon Alexa, Google Assistant, and Microsoft Cortana. JMIR Mhealth Uhealth 2021 Jan 11;9(1):e24045 [FREE Full text] [doi: 10.2196/24045] [Medline: 33427680]
9. de Melo T, Figueiredo CMS. Comparing news articles and Tweets about COVID-19 in Brazil: sentiment analysis and topic modeling approach. JMIR Public Health Surveill 2021 Feb 10;7(2):e24585 [FREE Full text] [doi: 10.2196/24585] [Medline: 33480853]
10. Sezgin E, Huang Y, Ramtekkar U, Lin S. Readiness for voice assistants to support healthcare delivery during a health crisis and pandemic. NPJ Digit Med 2020 Sep 16;3(1):122 [FREE Full text] [doi: 10.1038/s41746-020-00332-0] [Medline: 33015374]
11. Goh ASY, Wong LL, Yap KYL. Evaluation of COVID-19 information provided by digital voice assistants. International Journal of Digital Health 2021;1(1):3 [FREE Full text] [doi: 10.29337/ijdh.25]
12. Alagha EC, Helbing RR. Evaluating the quality of voice assistants' responses to consumer health questions about vaccines: an exploratory comparison of Alexa, Google Assistant and Siri. BMJ Health Care Inform 2019 Nov 21;26(1):j4008 [FREE Full text] [doi: 10.1136/bmjhci-2019-100075] [Medline: 31767629]
13. Pereira DA. A survey of sentiment analysis in the Portuguese language. Artif Intell Rev 2020 Jul 06;54(2):1087-1115. [doi: 10.1007/s10462-020-09870-1]
14. Coronavirus disease (COVID-19): Vaccines. World Health Organization. URL: https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-(covid-19)-vaccines [accessed 2021-07-31]
15. Krippendorff K. Content analysis: An introduction to its methodology. Thousand Oaks, CA: Sage Publications, Inc; 2018.
16. Fojas CL, Kim J, Minsky-Rowland JD, Algee-Hewitt BFB. Testing inter-observer reliability of the Transition Analysis aging method on the William M. Bass forensic skeletal collection. Am J Phys Anthropol 2018 Jan 26;165(1):183-193. [doi: 10.1002/ajpa.23342] [Medline: 29072307]

Abbreviations

CDC: Centers for Disease Control and Prevention
VA: voice assistant
WHO: World Health Organization

Edited by A Kushniruk; submitted 03.11.21; peer-reviewed by M Kopka, I Idrees; comments to author 28.11.21; revised version received 15.12.21; accepted 29.12.21; published 21.03.22.

Please cite as:
Seródio Figueiredo CM, de Melo T, Goes R
Evaluating Voice Assistants' Responses to COVID-19 Vaccination in Portuguese: Quality Assessment
JMIR Hum Factors 2022;9(1):e34674
URL: https://humanfactors.jmir.org/2022/1/e34674
doi:10.2196/34674
PMID:35041617

Original Paper

A Virtual Coach (Motibot) for Supporting Healthy Coping Strategies Among Adults With Diabetes: Proof-of-Concept Study

Giulia Bassi[1], PhD; Claudio Giuliano[2], PhD; Alessio Perinelli[3], PhD; Stefano Forti[2], PhD; Silvia Gabrielli[2], PhD; Silvia Salcuni[1], PhD

[1]Department of Developmental and Socialization Psychology, University of Padova, Padova, Italy

[2]Digital Health Lab, Centre for Digital Health and Wellbeing, Fondazione Bruno Kessler, Trento, Italy

[3]CIMeC, Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy

Corresponding Author:
Giulia Bassi, PhD
Department of Developmental and Socialization Psychology
University of Padova
Via Venezia 8
Padova, 35131
Italy
Phone: 39 3477334405
Email: giulia.bassi@phd.unipd.it

Abstract

Background: Motivation is a core component of diabetes self-management because it allows adults with diabetes mellitus (DM) to adhere to clinical recommendations. In this context, virtual coaches (VCs) have assumed a central role in supporting and treating common barriers related to adherence. However, most of them are mainly focused on medical and physical purposes, such as the monitoring of blood glucose levels or following a healthy diet.

Objective: This proof-of-concept study aims to evaluate the preliminary efficacy of a VC intervention for psychosocial support before and after the intervention and at follow-up. The intent of this VC is to motivate adults with type 1 DM and type 2 DM to adopt and cultivate healthy coping strategies to reduce symptoms of depression, anxiety, perceived stress, and diabetes-related emotional distress, while also improving their well-being.

Methods: A total of 13 Italian adults with DM (18-51 years) interacted with a VC, called Motibot (motivational bot) using the Telegram messaging app. The interaction covered 12 sessions, each lasting 10 to 20 minutes, during which the user could dialogue with the VC by inputting text or tapping an option on their smartphone screen. Motibot is developed within the transtheoretical model of change to deliver the most appropriate psychoeducational intervention based on the user's motivation to change.

Results: Results showed that over the 12 sessions, there were no significant changes before and after the intervention and at follow-up regarding psychosocial factors. However, most users showed a downward trend over the 3 time periods in depression and anxiety symptoms, thereby presenting good psychological well-being and no diabetes-related emotional distress. In addition, users felt motivated, involved, encouraged, emotionally understood, and stimulated by Motibot during the interaction. Indeed, the analyses of semistructured interviews, using a text mining approach, showed that most users reported a perceived reduction in anxiety, depression, and/or stress symptoms. Moreover, users indicated the usefulness of Motibot in supporting and motivating them to find a mindful moment for themselves and to reflect on their own emotions.

Conclusions: Motibot was well accepted by users, particularly because of the inclusion of mindfulness practices, which motivated them to adopt healthy coping skills. To this extent, Motibot provided psychosocial support for adults with DM, particularly for those with mild and moderate symptoms, whereas those with severe symptoms may benefit more from face-to-face psychotherapy.

(JMIR Hum Factors 2022;9(1):e32211) doi:10.2196/32211

KEYWORDS

virtual coach; diabetes mellitus; adults; psychosocial factors; mindfulness; proof-of-concept study; mobile phone

Introduction

Background

Physical, medical, and psychosocial factors significantly contribute to adherence rates to the clinical recommendations in adults with diabetes mellitus (DM) by promoting or hindering optimal diabetes self-management. Appropriate diabetes self-management is central to long-term diabetes care, and it includes several healthy behaviors, such as monitoring of glycemic levels, physical exercise, healthy eating, taking prescribed medication and/or insulin injections, which in turn have an impact on the general well-being of people with DM. However, these healthy behaviors are difficult to maintain. Indeed, studies have shown that high levels of diabetes-emotional distress are associated with a worsening of self-care behaviors as well as glycemic levels [1]. Diabetes-emotional distress is also a risk factor for stress, anxiety, and depression symptoms. Indeed, the prevalence rates of depression are much higher in people with DM than in the general population, in which they are estimated to be 17% [1]. With regard to anxiety symptoms, studies have found that 14% of adults with DM show generalized anxiety disorder, a prevalence much higher than the 3% to 4% rate identified in a community sample [2-4]. Anxiety is related to unhealthy lifestyle choices, such as augmented smoking prevalence, assumption of food high in cholesterol, and a sedentary lifestyle, which can all lead to poor disease management [5]. In addition, higher levels of anxiety hinder cognitive capacity, which in turn influences diabetes management and thus the ability to fully follow clinical recommendations [1,5]. Similarly, feeling stressed determines the release of stress hormones, such as cortisol and adrenaline, which prevent insulin from working properly (ie, insulin resistance) and thus increases glycemic levels [6]. Depression, anxiety, and stress are associated with the risk of developing cardiovascular diseases, and the presence of DM further increases this risk [1,5,6]. All together, these factors provoke lower adherence rates and impairment in the well-being of people with DM, leading to poor disease outcomes [7]. Therefore, one can assume the presence of a complex interplay between psychosocial factors and diabetes management, meaning that they influence one another. Thus, the American Association of Diabetes Educators (AADE) guidelines introduced a healthy coping construct to identify healthy coping strategies to reduce these symptoms and improve the general well-being of adults with DM [8]. In particular, AADE suggests strategies to cope with life stresses and the challenges of managing DM, such as meditating [8]. Indeed, several studies have demonstrated the efficacy of mindfulness practices in emphasizing self-acceptance in the general population [9] and treating depressive symptoms in people with DM [10,11]. In this context, motivation is a core component in adherence to the diabetes regimen as it is specifically conceptualized for its process rather than for a specific goal [12]. Indeed, the transtheoretical model of change (TTMC) [13] defines motivation as a continuum rather than as an all-or-nothing construct, in which the individual can move across 5 stages (ie, precontemplation, contemplation, preparation, action, and maintenance), thus moving forward or backward. In this regard, digital health technologies play a central role in promoting health care, especially in chronic diseases. For instance, TTMC has been widely used in digital solutions to predict or evaluate behavioral changes in physical activity [14], diet [15], and glycemic control [16] and to ameliorate adherence to medications in adults with risk factors for the onset of cardiovascular diseases [17,18].

Virtual Coaches: User Engagement and User Experience

The increasing rates of diabetes worldwide are a problem for the diabetologists treating diabetes, who already, as things stand, do not have enough time to assist every patient with physical, medical, and psychosocial issues. In this regard, virtual coaches (VCs) have become an increasingly relevant resource for the management of chronic diseases and for promoting behavioral changes in the self-management of individuals. Indeed, they aim to provide personalized guidance and improve intervention outcomes by mimicking human beings [19]. Indeed, VC in the health care field is mainly aimed at developing personalized user-system interactions and supporting individuals in their behavioral changes [20,21]. This is important for improving user engagement (UE) and compliance, both of which are crucial for achieving long-term behavioral changes and adjustment toward a healthier lifestyle [22]. UE is a multifaceted construct that refers to the quality of the user experience (UX), including the individual's time, cognitive, affective, and behavioral investment during the interaction with a digital solution [23]. The UE construct goes beyond user satisfaction: indeed, the literature suggests that the capacity to engage and maintain engagement in the interaction with a digital solution seems to show positive results in eHealth, e-learning, and web searching [23]. Indeed, prolonged engagement has been shown to be promising in a diabetes prevention program with the use of VC [24], engaging 69% of adults for the whole study and resulting in 8.98% weight loss [24]. In addition, in 2 other recent studies, adults with type 2 DM (T2DM) [25] and young adults with type 1 DM (T1DM) [26] reported feeling engaged and satisfied with a VC embedded in an application. In other studies, people with DM reported an increased level of satisfaction with the interaction with the VC [27,28]. Indeed, VCs for people with DM seem to favor self-care behaviors and behavioral changes as well as support them at follow-up. For instance, a recent review reported that VCs for people with DM represent effective interventions for fostering their glycemic control, in combination with standard care [29]. Therefore, VCs seem to be capable of overcoming common barriers related to adherence by delivering data-driven personalized support in real time and being available at any time during the day [30], thus allowing scalability. In this regard, UX is a crucial element that intersects the UE. Taking into consideration the definition proposed by the International Organization for Standardization [31], UX includes users' engagement, pleasure, desirability, values, emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors, and accomplishments, which occur before, during, and after the use of a digital solution. The International Organization for Standardization also lists 3 factors that influence UX: the user's current state and previous experience, system properties, and use context [31]. Therefore,

understanding users' needs, their working environments, interactions, and emotional reactions can help design VCs from the UX point of view [31]. Thus, the user becomes an active contributor to this process [31]. Studies have shown that higher levels of UX have been associated with an increased effectiveness of digital health interventions targeting improvements in T2DM self-management [25], physical activity [32], and diet [33]. However, there is a lack of evidence regarding UE and UX as constructs that interplay in the development, evaluation, and implementation of VCs for psychosocial support of adults with DM.

Comparison With Previous Work

To our knowledge, this is the first study to implement VC for psychosocial support in adults with T1DM and T2DM. Notwithstanding the originality of this work, in previous studies, VCs have been deployed to improve healthy coping strategies in college students, showing their beneficial effect in reducing symptoms of distress [34,35]. With regard to the development of VCs in the field of diabetes, studies have designed a conversational agent [36] and an interactive diary [37,38]. These digital interventions were both embedded in a smartphone app to improve health-related quality of life among adults with T2DM [35] and T1DM [37,38]. Health-related quality of life is an important and well-known construct that underlies the concept of general well-being. However, it should also be noted that anxiety, depression, and stress symptoms interplay with diabetes management. This means that these outcomes can hinder an individual's ability to manage diabetes and maintain effective glycemic control. Hence, it is important to include these variables when developing programs and interventions for adults with DM.

Objectives

Bearing all these aspects in mind, this VC (Motivational bot—Motibot) aims to support and motivate adults with T1DM and T2DM to adopt healthy coping strategies. In turn, these healthy coping strategies should reduce depression, anxiety, perceived stress symptoms, and diabetes-related emotional distress and improve well-being. Therefore, the aim of this proof-of-concept study was threefold:

1. To evaluate the preliminary efficacy of the VC intervention before and after the intervention and at follow-up in reducing the abovementioned psychosocial symptoms while also improving the well-being of adults.
2. To investigate UX and UE with the VC for psychosocial support accessed through personal smartphones within the Telegram messaging app.
3. To evaluate semistructured interviews on both UX and how users felt during their interaction with the VC.

Methods

Participants and Recruitment

The study involved 18 voluntary adults with T1DM and T2DM recruited in Italy via social network sites (ie, Facebook groups) using snowball sampling. Five adults dropped out of the study for personal and medical reasons. Therefore, the final sample included 13 adults aged between 18 and 51 years (mean 30.08, SD 10.61 years), 77% (10/13) of which were women; 62% (8/13) of adults had T1DM and 39% (5/13) had T2DM, with an overall mean diabetes duration of 10 (SD 8.49) years. One participant did not complete the psychological measures after the intervention and was therefore excluded from the analyses of the psychosocial variables. The inclusion criteria for participating in this study were as follows: (1) having T1DM or T2DM and (2) owning a smartphone and a Telegram account. The decision to include both types of DM was guided by the notion that there are similarities between the lifestyle guidelines for adults with T1DM and T2DM, as emerged from the results of a recent meta-analysis [39]. Participants were excluded if they had gestational diabetes or prediabetes.

Procedure and Ethics

This work is a proof-of-concept study for adults with T1DM and T2DM, conducted following the Obesity-Related Behavioral Intervention Trials (ORBIT) framework [40]. The ORBIT model supported guidance throughout the whole process as it emphasizes the importance of adopting a data-driven iterative approach to optimize subsequent iterations of the intervention [40]. In particular, this model places the user at the center of the design process. The study was conducted in compliance with the Declaration of Helsinki (Italian law 196/2003, European Union General Data Protection Regulation 679/2016). The Interdepartmental Ethical Committee of Psychology of the University of Padova (Italy) approved the project (approval number: 3968; February 3, 2021), stating that there were no critical ethical issues. The participants signed a written informed consent sent via mail, agreeing to participate in the study and semistructured interviews 1 month after the end of the study. They were informed that their data would be confidential, that they could omit any information they did not wish to give, and that they could withdraw from the study at any moment without having to provide any explanation.

Intervention Description: Motibot Design

Motibot (Figure 1) is a VC designed to provide psychosocial support by motivating adults with DM to adopt and cultivate healthy coping strategies, which, according to the AADE guidelines, should be flexible and adaptable to the users' needs [8]. These coping strategies, in turn, should foster adults' well-being by reducing depression, anxiety, and perceived stress symptoms and diabetes-related emotional distress. Motibot was developed by the Digital Health Lab at Fondazione Bruno Kessler Research Center for Digital Health and Wellbeing using Rasa [41], an open-source platform designed for the development and training of VCs. It was then deployed through the Telegram messaging app. The environment provided by Rasa exploits machine learning (ML) libraries and pretrained embeddings from language models, thus allowing the construction of a VC for a specific language by combining ML approaches and handcrafted rules. Motibot relies on natural language understanding (NLU) [42], which is an ML technique that enables the VC to interpret user messages. NLU, together with the conversational history and a set of predefined variables, determines the transition from one turn of the dialogue to another. In this study, the NLU system was trained by feeding it with a data set comprising 6899 examples of user utterances

categorized by intent and annotated with entities. Examples of intents are to *affirm, deny, say your name, say what you feel, schedule the next meeting,* and *xpress the level of motivatio*ne. Examples of entities are the *user's name, the emotion felt, the date, time of the next meeting*, and *level of motivation*. NLU was used to interpret the intents and entities. In this study, we defined 54 intents and 6 entities. During interactions between Motibot and users, intents and entities were extracted from users' messages and classified using a trained multitask transformer architecture. Motibot was designed to last for 12 sessions of 10-20 minutes each, during which Motibot interacted with the user according to the scripts previously defined as shown in Figure 2. Because of the flexibility of this VC, specifically designed to be as adaptable as possible to the users' daily life, every session was initiated following a scheduled plan decided by the users themselves to best suit their needs. Users were able to respond to Motibot by inputting text or tapping an option on their smartphone screen. The interaction between Motibot and users is designed considering evidence-based approaches related to counseling and psychoeducation as displayed in Figure 2. In particular, these approaches are linked to the healthy coping construct [8] and to mindfulness-based cognitive therapy [43] to support and motivate the development and/or enhancement of coping strategies. For these reasons, the whole conversational protocol was developed referring to TTMC [13], which allows the VC to understand what motivational state the user is in and consequently deliver the most appropriate psychoeducation intervention, which is based on the user's motivation to change. At the beginning of the first session, Motibot asks the users to present themselves by telling they are. Subsequently, Motibot delivers a video presentation of itself, its functionality, and its main features to involve the user in the interaction. Thereupon, Motibot delivers 3 different questionnaires to assess the levels of depression (Patient Health Questionnaire-9 [PHQ-9]) [44], anxiety (Generalized Anxiety Disorder-7 [GAD-7]) [45], and perceived stress symptoms (Perceived Stress Scale-10 [PSS-10]) [46]. These 3 questionnaires were also sent after the intervention and at follow-up. In this last case, 2 psychosocial scales were added to assess diabetes-related emotional distress (Problem Areas in Diabetes Scale–Short Form-5 [PAID-5]) [47] and general well-being (World Health Organization-5 Well-Being Index [WHO-5]) [48]. PAID-5 and WHO-5 were evaluated only at follow-up, that is, 2 months after the end of the study. The latter 2 scales were included to verify whether coping strategies had been internalized, thus leading to greater well-being. Indeed, diabetes self-management is well known to be influenced by these outcomes. In addition, Motibot sent 2 other questionnaires to assess UX during the whole interaction and UE only at the end of the intervention to comprehend the users' overall and final involvement. One month after the end of the study, semistructured interviews were conducted to understand both UX and how users felt during their interaction with Motibot.

Figure 1. Motibot: the virtual coach.

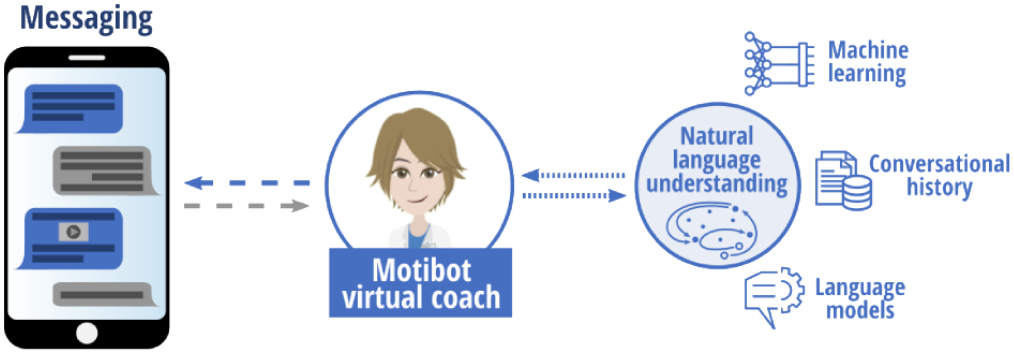
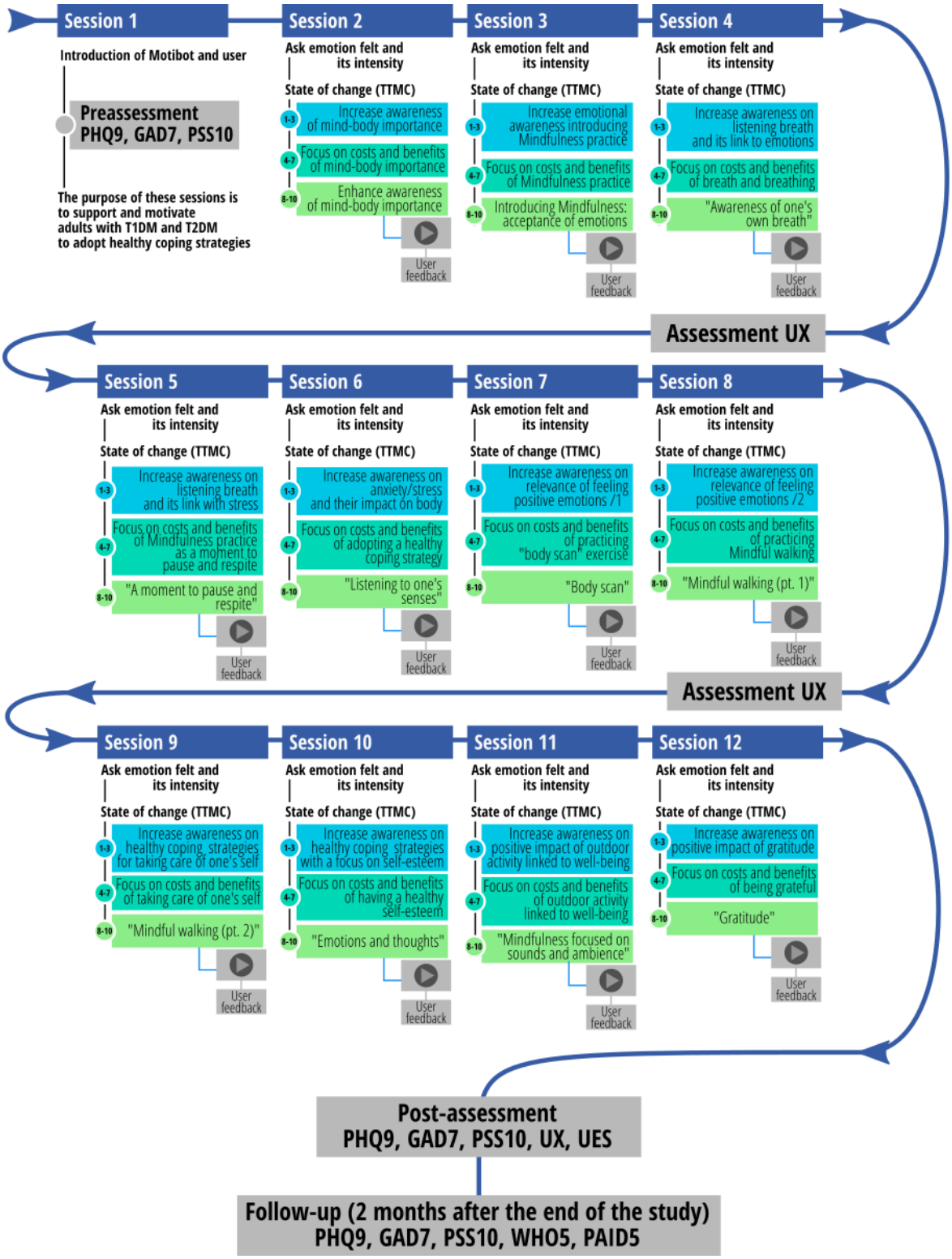


Figure 2. Graphical representation of the conversational protocol delivered to users and its chronological structure. GAD7: Generalized Anxiety Disorder-7 Items; PAID5: Problem Areas in Diabetes Scale–Short Form-5 Items; PHQ9: Patient Health Questionnaire-9 Items; PSS10: Perceived Stress Scale-10 Items; T1DM: type 1 diabetes mellitus; T2DM: type 2 diabetes mellitus; TTMC: transtheoretical model of change; UES: User Engagement Scale; UX: user experience; WHO5: The World Health Organization-5 Well-Being Index.



Each day, Motibot asks the users what emotion they are feeling at that precise moment as well as its intensity, to support them in becoming more aware of their own emotions and self-reflecting on them. After this question, according to TTMC and following the state of change ruler (ie, precontemplation state, contemplation, preparation, action, and maintenance) [13], Motibot asks users "How much do you want to improve your well-being on a scale from 1 (not at all) to 10 (very much)?" to understand their motivation to maintain diabetes under control. When users are considered in the *precontemplation state*, Motibot tries to investigate why they feel this way and then attempts to increase emotional awareness by helping them to self-reflect on their emotions and on the importance of taking care of both their body and mind. When users are considered

in the *contemplation state*, Motibot provides motivational interventions, which focus the attention on the costs and benefits of adopting a healthier behavior to favor psychosocial well-being. Finally, when users are in the *action state*, Motibot provides behavioral interventions by sending audio tracks related to mindfulness practices.

Data Collection

As reported in Textbox 1, the psychosocial questionnaires were administered before and after the intervention and at follow-up. Moreover, semistructured interviews were conducted 1 month after the end of the study.

Textbox 1. Overview of questionnaires and of their administration timing.

- Before the intervention
  - Patient Health Questionnaire-9 (depression)
  - Generalized Anxiety Disorder-7 (anxiety)
  - Perceived Stress Scale-10 (perceived stress)
- At the 2nd, 8th, and 12th sessions
  - User Experience Questionnaire (user experience)
- After the intervention
  - Patient Health Questionnaire-9
  - Generalized Anxiety Disorder-7
  - Perceived Stress Scale-10
  - User Engagement Scale–Short Form (user engagement)
- 1 month after the end of the study
  - Semistructured interviews
- At follow-up (2 months after the end of the study)
  - Patient Health Questionnaire-9
  - Generalized Anxiety Disorder-7
  - Perceived Stress Scale-10
  - World Health Organization-5 Well-Being Index (well-being)
  - Problem Areas in Diabetes Scale–Short Form-5 (diabetes-related emotional distress)

The *PHQ-9* [44] is a brief self-reported unidimensional measure developed to assess and monitor the severity of depression symptoms in the previous 2 weeks. The questionnaire includes 9 items rated on a 4-point Likert scale (from 0 *never* to 3 *almost every day*). The PHQ-9, which incorporates the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision criteria, has a total score ranging from 0 to 27, with a score of 10 representing the optimal cutoff to detect clinically relevant depression. The PHQ-9 comprises five categories of severity: (1) absent (scores 0-4), (2) subthreshold depression (scores 5-9), (3) mild depression (scores 10-14), (4) moderate depression (scores 15-19), and (5) major depression (scores 20-27). An example of an item is the following: “During the last two weeks, on how many days did you feel little interest or pleasure in doing things?” (item 1). The PHQ-9 has shown good psychometric properties [44]. The *GAD-7* [45] is a brief self-reported unidimensional measure aimed at screening probable cases of GAD and assessing the severity of symptoms in the previous 2 weeks. The questionnaire comprises 7 items based on a 4-point Likert scale (from 0 *never* to 3 *almost every day*). The GAD-7 incorporates the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision criteria and has a total score ranging from 0 to 21, with a score of 10 as the cutoff for GAD. The questionnaire included three categories of severity: (1) mild anxiety symptoms (score ≥5), (2) moderate anxiety symptoms (score ≥10), and (3) severe anxiety symptoms (score ≥15). An example of an item is the following: “In the last two weeks, how often did each of the following problems bother you? Feeling nervous, anxious, or tense” (item 1). The GAD-7 has demonstrated good validity and reliability [45]. The *PSS-10* [46] is a brief self-reported unidimensional measure that assesses an individual’s perception of stress in the previous month. The PSS is a measure of the degree to which each situation in one’s life is perceived as stressful; indeed, the items are designed to evaluate the degree to which individuals find their lives unpredictable, uncontrollable, or overloaded. The scale also contains a series of direct questions about the current levels of perceived stress. The PSS consists of 10 items based on a 4-point Likert scale (from 0 never to 5 very often). The total PSS score ranges from 0 to 40, with high scores indicating a high level of perceived stress. PSS includes three categories of severity: (1) low perception of stress (scores 0-13), (2) moderate perception of stress (scores 14-26), and (3) high perception of stress (scores

27-40). An example of an item is the following: “In the last month, how often have you felt out of sorts because something unexpected happened?” (item 1). The PSS-10 has demonstrated good psychometric properties regarding reliability and validity [46]. The *PAID-SF-5* [47] is a self-reported unidimensional measure aimed at assessing diabetes-related emotional distress. The questionnaire comprises 5 items based on a 5-point Likert scale (from 0 *not a problem* to 4 *serious problem*). Total scores range from 0 to 100, with higher scores (ie, ≥40) indicating greater diabetes-related emotional distress. The PAID-SF-5 has demonstrated good psychometric properties [47]. The *WHO-5* [48] is a self-reported unidimensional measure that evaluates psychological well-being, a core dimension of quality of life. The questionnaire comprises 5 items rated on a 6-point Likert scale (from 0 never to 5 always). The total score was rescaled to range between 0 and 100, with a score ≤50 suggesting poor psychological well-being and a score ≤28, indicating depression, showing good psychometric properties [48]. The *User Engagement Scale–Short Form* (UES-SF) [49] is a brief self-report questionnaire aimed at assessing user engagement with a digital solution. The UES-SF includes 12 items based on a 5-point Likert scale (from 1 *strongly disagree* to 5 *strongly agree*). The UES-SF comprises 4 factors: (1) focused attention, which indicates the feeling of being immersed in the interaction (eg, “I lost myself in this experience”); (2) perceived usability, which is the negative affect experienced owing to the interaction and the effort spent (eg, “I felt frustrating while using Motibot”)—this factor is the only one in which the scores were reversed; (3) aesthetic appeal, which represents the graphical and visual appeal related to the digital solution (eg, “Motibot was aesthetically appealing”); and (4) the reward factor (eg, “Using Motibot was worthwhile”). The latter is a single set of 3 factors related to the original UES questionnaire [49,50], such as the endurability, which evaluates the overall success of the interaction; the novelty, which examines the overall interest related to the interaction with a digital solution; and finally, the felt involvement factor, which evaluates the overall fun interaction. The overall scale was found to be reliable [49]. The *User Experience Questionnaire* (UEQ) used in this study is an adapted version of the original UEQ [51], modified ad hoc to make the bipolar adjectives more appropriate to the aims of this study. In particular, the questionnaire included 28 adjectives, either positive or negative, designed to assess the experience of interacting with the VC. Each item was scored on a 5-point Likert scale (from 1 *strongly disagree* to 5 *strongly agree*). Textbox 2 shows the selection of items for this study.

Textbox 2. The items of the User Experience Questionnaire.

Positive items

- Pleasant
- Profound
- Cordial
- Comprehensible language
- Empathetic
- Attentive
- Motivating
- Encouraging
- Supportive
- Trustworthy
- Flexible
- Interesting
- Effective

Negative items

- Annoying
- Not reliable
- Unappealing
- Unclear
- Complicated
- Not efficient
- Too much information
- Dissuading
- Not stimulating
- Not engaging
- Unpredictable
- Not reflective
- Conventional
- Not effective
- Rigid

Semistructured interviews were conducted by GB with all participants who concluded the interaction with Motibot. The interviews were based on 11 ad hoc questions administered 1 month after the end of the study and lasted approximately 10 minutes. Each interview started with asking the motivation for participating in this study and concluded with a question in which the participant should explain whether they would suggest Motibot to other people with the same chronic illness, explaining the reason. The other 9 questions were divided into 2 sections as reported in Textbox 3. The first included 5 questions related to the experience that users had with Motibot; therefore, the goal was to assess the UX. On the other hand, the second section included 4 questions related to how users felt during the interaction with Motibot from a psychological perspective.

Textbox 3. Questions asked to participants during semistructured interviews.

- What motivated you to participate in the study?
  - User experience
    - What were your expectations with regard to Motibot?
    - With regard to Motbiot, which aspect did you like the most?
    - With regard to motibot, which aspect did you dislike the most?
    - With regard to Motibot, how was your user experience?
    - Would you be interested in using, in the future, a complete virtual coach?
- How users felt during the interaction
  - Motibot proposed to you several audio tracks regarding mindfulness. How did you live them?
  - Did you find Motibot useful to find a mindful moment for yourself?
  - Did Motibot help you to soothe any anxiety, stress and/or depression symptoms?
  - Did you listen to Motibot mindfulness audio tracks again at the end of the study?
- Would you suggest Motibot to someone with diabetes mellitus? Why?

Statistical Analysis

Statistical analyses were performed using R, version 4.0.0 (The R Foundation for Statistical Computing) [52], and SPSS Statistics, version 24.0 (IBM Corp) [53]. The Shapiro-Wilk test was performed to evaluate the normality of the sample distributions of the variables investigated in this study. Descriptive analysis was carried out on psychological dimensions, namely depression, anxiety, and perceived stress, before and after the intervention and at follow-up. The same analysis was performed on diabetes-related emotional distress and well-being, although only at follow-up. All data are shown as plots. The Kruskal-Wallis nonparametric test was used to evaluate differences in depression, anxiety, and stress among participants. A post hoc Wilcoxon nonparametric test was performed to compare the differences in the aforementioned outcomes before and after the intervention and at follow-up to understand whether the psychoeducational intervention had been effective. Means and SDs were computed for UX, which was evaluated at the 4th, 8th, and 12th sessions and for UE, which was evaluated at the end of the study. The data regarding UX are displayed as plots. A text mining approach [54,55] was followed to extract information from the semistructured interviews on UX and on how users felt during the interaction with Motibot. This analysis was implemented by relying on the *Quanteda* R package [56] and on custom shell scripting code under a Linux environment. The analysis was carried out on the written interview transcripts (in Italian) as follows: first, they were cleaned by replacing uppercase letters and removing numbers, punctuation, and stopwords. Thereupon, user's answers were divided into groups, each containing all answers to one of the interview questions. Two analyses steps were implemented: (1) extraction, from some of the questions in the semistructured interviews, of 3 sets of responses (ie, yes/no/maybe) and (2) extraction, from the remaining questions, of recurrent concepts (ie, word stems) and their relations in terms of digrams (ie, pairs of word stems). A word stem was deemed to be recurrent if it appeared at least three times across interviews, whereas digrams were considered recurrent if they appeared at least two times. The criterion of 3 occurrences as the threshold for including a stem was chosen according to the following rule of thumb. We assume stems to be significant if they belong to the 5% most recurrent ones. However, because occurrence is quantified by an integer number, this percentile threshold can be enforced only approximately. Setting the criterion of minimum occurrences to 3 yielded the extraction, for the different questions, between 3.8% and 7.9% most recurrent stems (average 6.2%), in reasonable compliance with the 5% threshold assumed above. In addition, the average occurrence of stems for a given question was 1.35; a threshold of 3 occurrences was equivalent to the requirement of a stem recurring more than twice as frequently as the average.

Results

Preliminary Analysis

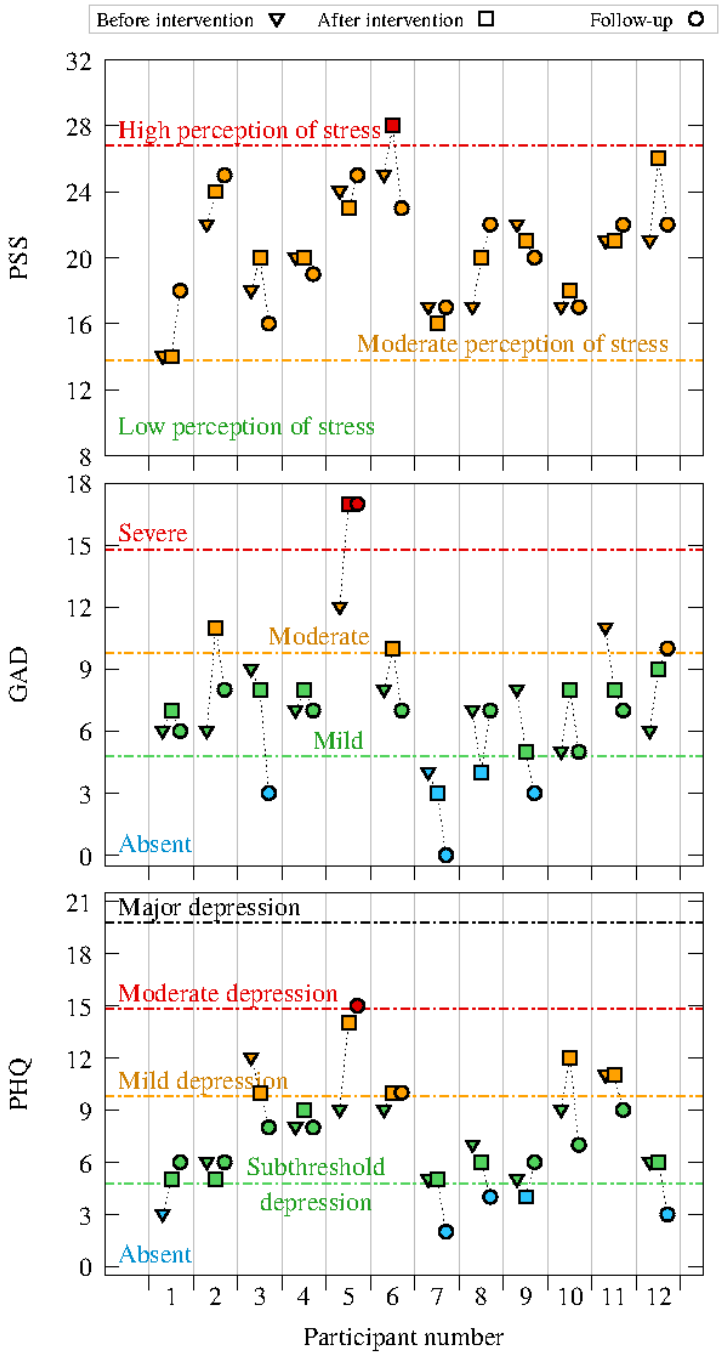
The variables investigated in this study showed a nonnormal distribution. No missing data were identified, and each participant answered all the questions administered. As mentioned in the *Participant* section, only one participant did not answer the entire questionnaire sent after the intervention and thus was excluded from the analyses regarding the psychosocial variables.

Perceived Stress, Anxiety, and Depression Symptoms

Overall, as displayed in Figure 3, all participants showed moderate symptoms concerning perceived stress (assessed using the PSS-10). Participant 6 was the only participant showing a high perception of stress after the intervention; however, the level of perceived stress diminished at follow-up. Data concerning anxiety and depression symptoms (assessed respectively through the GAD-7 and PHQ-9) seem to increase after the intervention and decrease at follow-up, except for participant 5, who presented severe symptoms at both time

points, thereby resulting in an outlier. The presence of an outliers did not affect the overall trend of data regarding these outcomes.

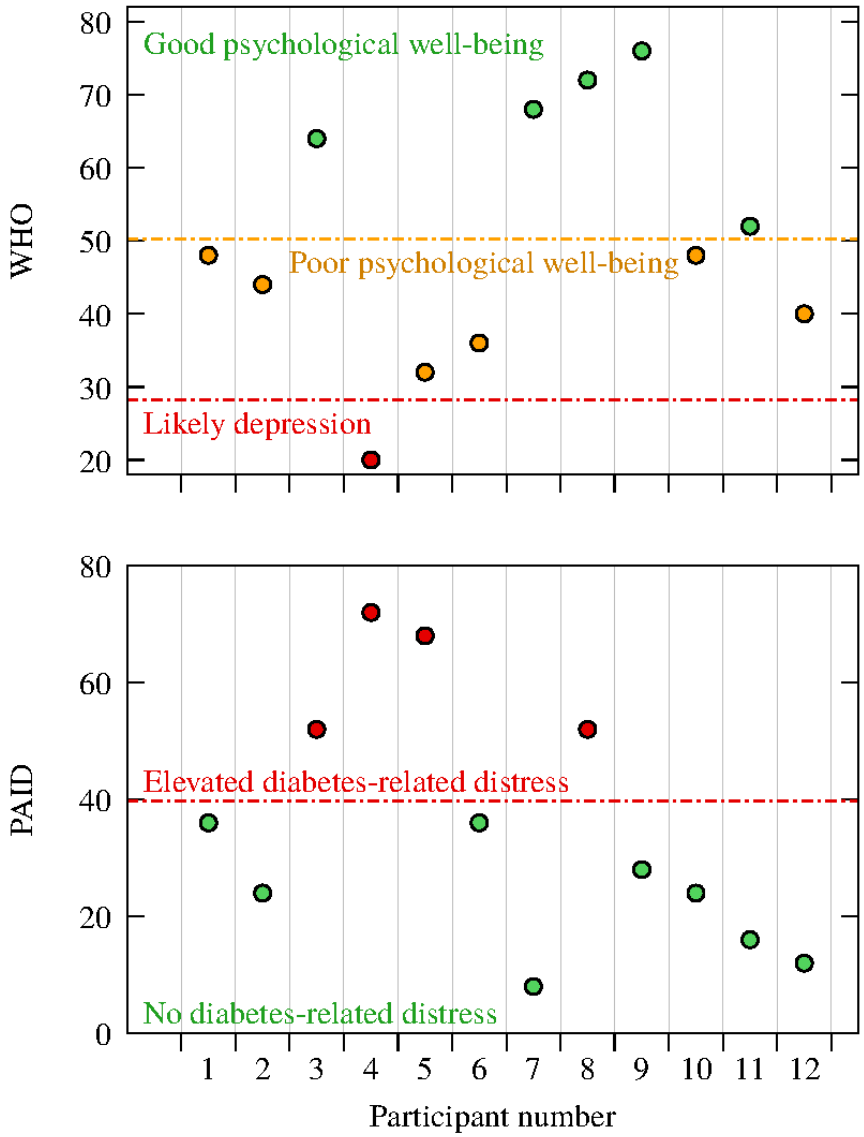
Figure 3. Plots of perceived stress, anxiety, and depression symptoms, assessed through the Perceived Stress Scale-10 (PSS-10), Generalized Anxiety Disorder-7 (GAD-7), and Patient Health Questionnaire-9 (PHQ-9), respectively, before and after intervention and at follow-up (N=12).



Well-being and Diabetes-Related Emotional Distress

As shown in Figure 4, when considering the presence of an outlier, participants ranged between poor and good psychological well-being (assessed through the WHO-5), with an overall mean of 50.00 (SD 17.18), which indicates an overall poor psychological well-being. However, if the outlier is excluded, the overall mean is 51.64 (SD 17.01), which corresponds to an overall good psychological well-being. With regard to diabetes-related emotional distress (assessed through the PAID-5), most participants did not present diabetes-related emotional distress; indeed, the overall mean was 35.67 (SD 21.20). If the approach described above is applied and therefore if the outlier is excluded, the overall mean is 32.73 (SD 19.50), which, being an even smaller value, suggests low levels of diabetes-related emotional distress.

Figure 4. Plots of well-being and diabetes-related emotional distress, evaluated through the World Health Organization-5 Well-Being Index (WHO-5) and Problem Areas in Diabetes Scale-5 (PAID-5), respectively, at follow-up (N=12).



Kruskal-Wallis Test for Psychosocial Outcomes

The Kruskal-Wallis test, carried out to assess the differences among depression, anxiety, and perceived stress symptoms, did not yield any significant results when considering the intervention period. Nevertheless, as shown in Figure 3, a downward trend can be identified over the 3 time periods (ie, before the intervention, after the intervention, and during follow-up).

User Experience

Analyses regarding the positive items (assessed through UEQ) reported a mean >3 on a 5-point Likert scale (mean 4.04, SD 0.22). In particular, the items *comprehensible language*, *empathetic*, *motivating*, *encouraging*, and *interesting* increased from the 2nd to the 12th session, whereas the item *supportive* tended to decrease from the 2nd to the 12th session as displayed in Figure 5. The specific means and SDs are reported in the Multimedia Appendix 1 (Table S1).

Analyses of the negative items (evaluated using UEQ), shown in Figure 6, reported a mean <2 on a 5-point Likert scale (mean 1.86, SD 0.30), thereby attesting that users disagreed with the overall items. In particular, the items *not stimulating*, *not engaging*, and *rigid* decreased from the 2nd to the 12th session. The specific means and SDs are reported in the Multimedia Appendix 1 (Table S2).

Figure 5. Plot of the positive items of the User Experience Questionnaire. Circled dots and error bars correspond to sample means and sample SDs, respectively (N=13).

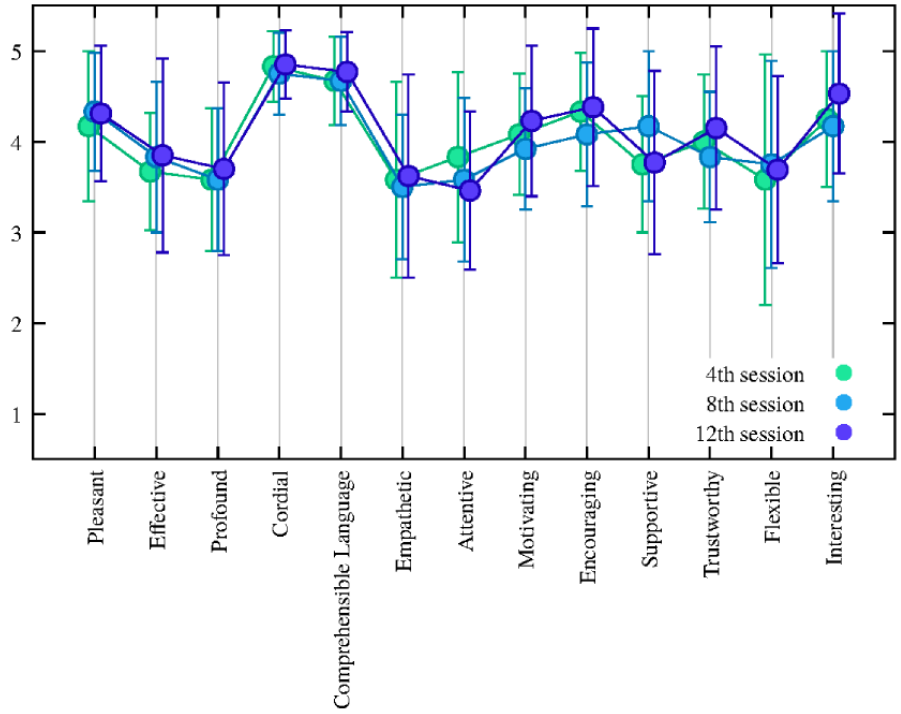
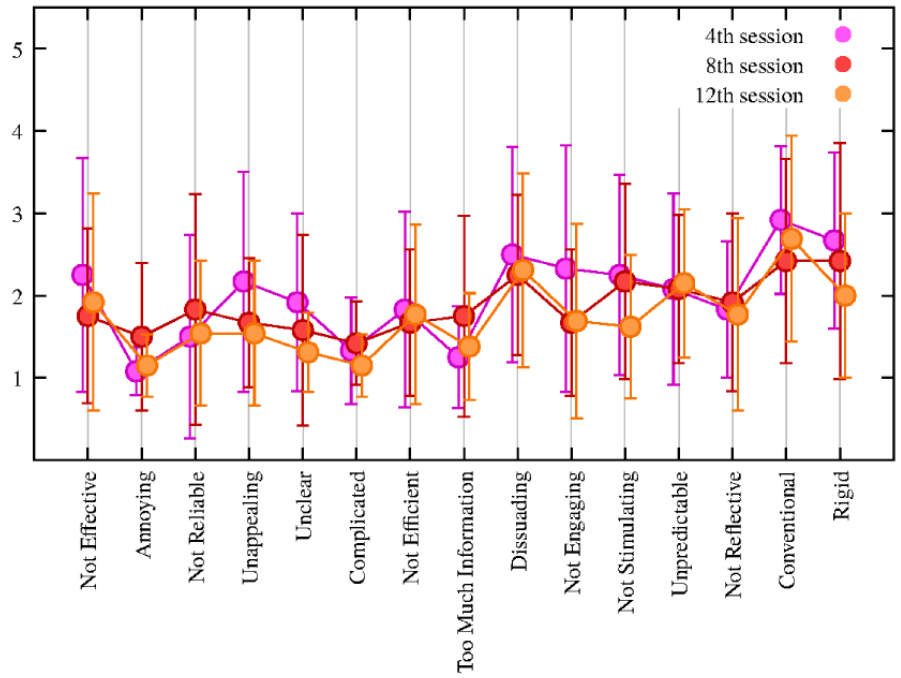


Figure 6. Plot of the negative items of the User Experience Questionnaire. Circled dots and error bars correspond to sample means and sample SDs, respectively (N=13).



User Engagement

Overall, the data on UE show that participants were engaged with Motibot, as reported in Table 1. The reward factor, which refers to the worthwhile and absorbing experience of the user with the digital solution, presents a maximum value of 5. The same result also emerged for perceived usability and focused attention. Notably, the perceived usability factor is the only factor in which the items were reversed, indicating a good effect experienced by the digital solution.

Table 1. Descriptive statistics for User Engagement Scale (N=13).

Parameters	Value, range	Value, mean (SD)
Total scale	3.25-4.83	4.14 (0.49)
Perceived usability	4-5	4.82 (0.32)
Focused attention	2.33-5	3.62 (0.83)
Esthetic appeal	2.67-4.67	3.79 (0.55)
Reward factor	3-5	4.33 (0.58)

Minimum and Maximum Scores of Participants Based on the User Engagement Questionnaire's Likert Scale

Text Mining

Overall, the average duration of the 13 semistructured interviews was 9.04 minutes. The transcripts of the interview answers comprised 562 words on average. The results of text mining applied to the answers of the semistructured interviews, concerning UX and how users felt during the interaction with Motibot, are graphically summarized in Figures 7 and 8, respectively. In both figures, bar plots show the distribution of 3 types of answers (ie, yes/no/maybe), whereas scatter plots highlight the most frequent concepts, namely word stems appearing at least three times in the interviews. Within scatter plots, arrows identify recurrent digrams, that is, sequences of 2-word stems appearing at least twice within the interviews. It is worth mentioning that the apparently opposite ordering of some digrams (eg, *support→psychological*) is because of the analysis being carried out on texts in Italian, in which word ordering is different from English. Stems were translated at the end of the analysis, considering the abovementioned potential nuances between the 2 languages. The radius of the circle is proportional to the number of occurrences of each stem.

Figure 7. Answers to the semistructured interviews related to user experience.

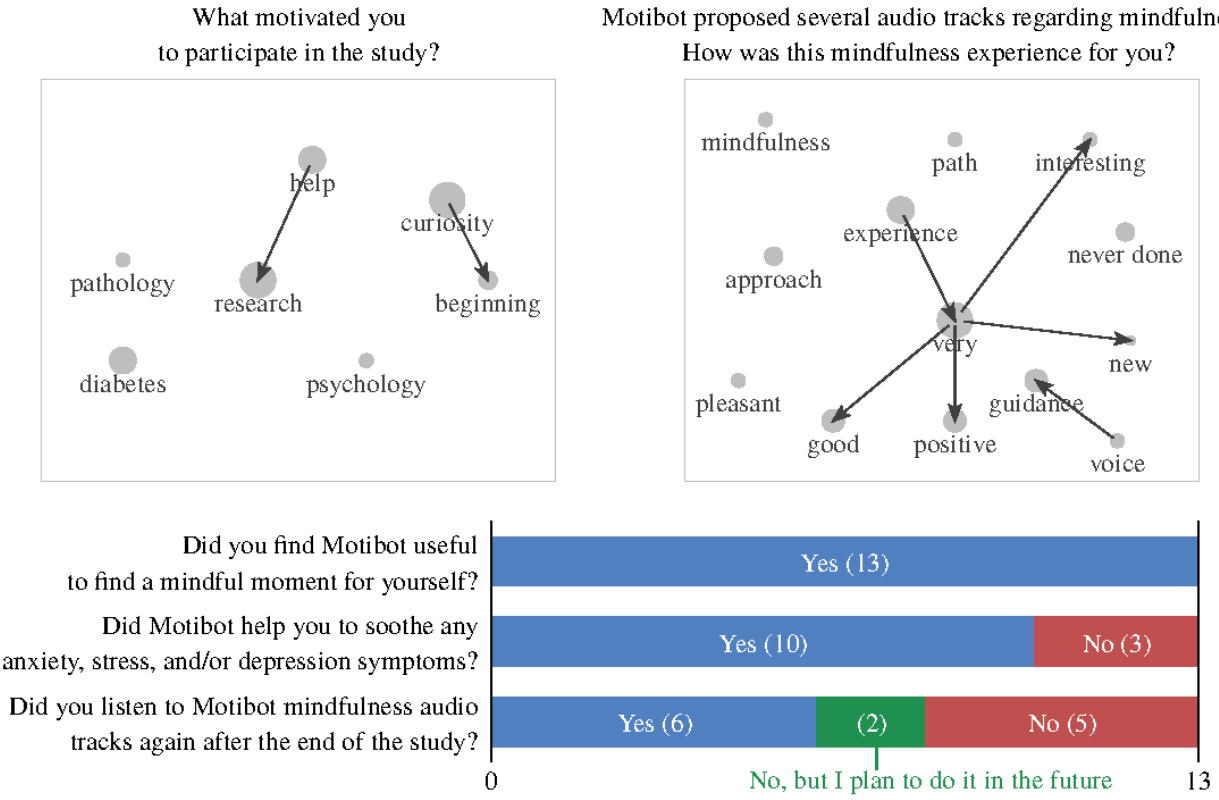
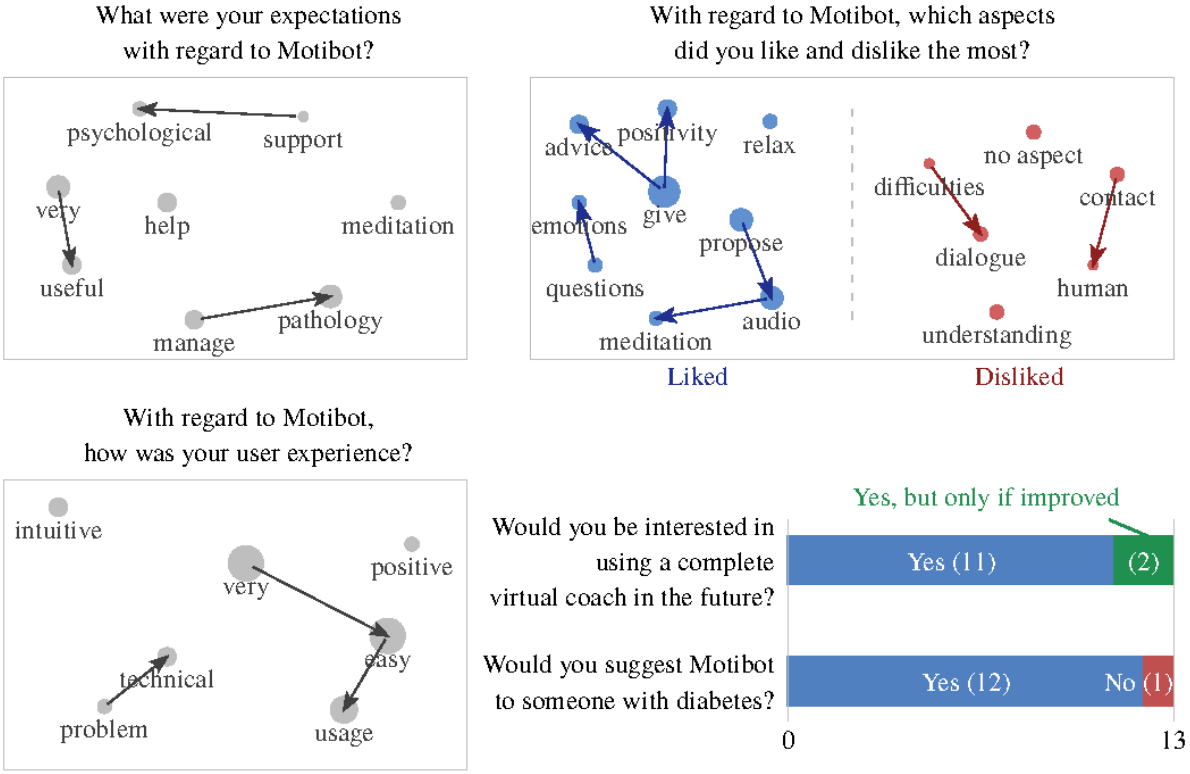


Figure 8. Answers to the semistructured interviews related to how users felt during the interaction with Motibot.



Text Mining: UX With Motibot

As shown in Figure 7, 85% (11/13) of the participants would be interested in using a VC for psychosocial support and 92% (12/13) would suggest VC to other people with the same chronic disease. Overall, participants reported having a positive experience with Motibot. As displayed in the upper-right panel, where stems graphed in blue and red correspond to *liked* and *disliked* aspects, respectively, users largely reported positive aspects in the interaction with Motibot, except for some technical problems. It should be mentioned that 3 participants reported that there were no aspects that they disliked.

Text Mining: How Users Felt During the Interaction With Motibot?

As displayed in Figure 8, users reported mindfulness audio tracks as a very good, positive, interesting, and new experience. They further considered the voice of the audio tracks as a guide to the mindfulness pathway. Indeed, 62% (8/13) of the participants also listened to the mindfulness audio again after the end of the study or planned to do so in the future, to grant themselves a further mindful moment. Finally, 77% (10/13) of the participants reported that Motibot helped them reduce the symptoms of anxiety, depression, and/or stress.

Discussion

Principal Findings

This proof-of-concept study evaluated the preliminary efficacy of a VC intervention for psychosocial support. Motibot, indeed, aims to support and motivate adults with T1DM and T2DM to adopt healthy coping strategies. In turn, these healthy coping strategies should reduce depression, anxiety, perceived stress symptoms, and diabetes-related emotional distress and improve well-being. This study further aims to evaluate UX and UE in the interaction with a VC from both qualitative and quantitative perspectives. Overall, preliminary evidence suggests that the digital intervention led to improvements in symptoms of anxiety and depression, as emerged from the downward trend of the related factors detected over the 3 time periods (ie, before the intervention, after the intervention, and during follow-up). Notably, participants showed an increase in anxiety and depression symptoms after the intervention and a subsequent decrease during follow-up, showing good psychological well-being, upon exclusion of an outlier, and no diabetes-related emotional distress. These data highlight how the effects of the psychoeducational intervention were maintained over time, thus leading to the users' internalization of healthy coping strategies and self-reflection of their own emotions. With regard to perceived stress, users showed moderate symptoms, which remained throughout the intervention, including follow-up. However, users did not present any diabetes-related emotional distress. These findings shed light on the possible stressful events that underpin perceived stress, such as the impact of contingent events; thus, the findings did not relate to the burden of managing DM. It is worth mentioning that this study was carried out in 2021 during the COVID-19 pandemic, which had a significant impact on psychological well-being among the whole population [57]. Overall, the users perceived a decrease in any symptoms of anxiety, depression, and/or stress, reporting the usefulness of Motibot in supporting and motivating them to find a mindful moment for themselves. The psychoeducational intervention was well accepted by users, particularly in the presence of a mindfulness *pathway*. Indeed, users reported a

very good, *positive*, *interesting*, and *new experience*: most users listened to the audio tracks even at the end of the study to achieve a mindful moment for themselves once again. Indeed, mindfulness-based interventions have recently become more relevant in the context of DM care, as they are associated with a reduction of negative emotions and an enhancement of an individual's attitude and coping strategies [58]. Users had a *positive* and *interesting experience* with Motibot, particularly because it proposed audio tracks relating to meditation and asked them what emotion they were feeling in that precise moment. The purpose of asking users to express their emotions is to motivate them to become more aware and reflect on them: one might expect that the more one is aware of their own emotions, the better they can regulate them. Motibot was perceived as *empathetic* and *stimulating* in its dialogic interaction, even if it was slightly less *supportive* from the 2nd to the 12th session. This last result might indicate that users become familiar with Motibot throughout the sessions and thus do not perceive any further support, albeit still feeling involved and absorbed in the interaction. However, Motibot was also perceived as *motivating* and *encouraging* in the adoption of healthy coping strategies: users appreciated that Motibot gave them *positive advice*. Notwithstanding these promising results, few users still reported a desire for *human contact* to receive psychological support. These data emerged particularly for those who presented with high levels of anxiety, depression, and/or perceived stress symptoms. Therefore, we speculate that VCs may be successfully used to support and motivate people with mild and moderate psychological symptoms, whereas those with more severe psychological symptoms may benefit more from psychotherapy support in face-to-face spontaneous and human settings. Furthermore, users encountered technical problems when interacting with Motibot, particularly when arranging the next session. However, this issue was addressed in the study. Nonetheless, users felt *involved* and *engaged* with Motibot, reporting a worthwhile and absorbing experience and a *positive* perception of use, stating that Motibot was *very easy to use*.

Limitations and Future Work

This study has 2 main limitations. First, the small sample size, which was chosen following the proof-of-concept phase related to the ORBIT model, as it allows the inclusion of few participants during the first phases of the design, evaluation, and implementation process [40]. However, this choice does not permit data generalization. Second, a more complex analysis approach concerning text mining on the semistructured interviews, such as supervised or unsupervised learning, could not be implemented owing to the relatively small number of participants and the limited length of the interviews (ie, between approximately 200 and 1000 words each). Future studies should integrate, in the development of a VC, medical factors such as the glycemic levels alongside the main psychosocial aspects, as they interplay with the management of DM and thus are variables worth analyzing. Finally, our future goal is to test Motibot with a larger sample size in a randomized controlled trial to investigate the effectiveness of the psychoeducational intervention in a systematic and controlled manner.

Conclusions

Motibot was developed through a combination of NLU and handcrafted rules with the aim of delivering a psychoeducational intervention for adults with T1DM and T2DM, which allows them to interact by using both free text and structured dialogue interaction. The results of this study showed positive user experience and engagement. In addition, the findings highlighted the usefulness of interacting with a VC to motivate adults with DM to adopt healthy coping strategies. These coping strategies, specifically related to mindfulness practices, allowed a reduction in anxiety, depression, and diabetes-related emotional distress symptoms, while also improving their well-being. This decrease in psychosocial symptoms and increase in well-being was also maintained at follow-up. VCs have the advantage of scalability, which leads to greater user accessibility, and thus, it is available at any time. Moreover, VCs are deployable to adults with DM who show mild and moderate psychosocial symptoms. In particular, VCs can provide them with valuable support, in combination with a dedicated psychotherapist both in a traditional face-to-face setting or in a digital solution referring to the stepped care model [59].

Acknowledgments

The authors would like to specially thank Gaia Dell'Arciprete for her support in proofreading this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mean and SD of the User Experience Questionnaire.
[DOCX File , 22 KB - humanfactors_v9i1e32211_app1.docx]

References

1. August K, Kelly C, Abbamonte J. Psychosocial factors associated with diabetes self-management. In: Glucose Intake and Utilization in Pre-Diabetes and Diabetes. Amsterdam, Netherlands: Elsevier; 2014.

2. Generalized anxiety disorder. In: Psychiatric Disorders in America: The Epidemiological Catchment Area Study. New York: The Free Press; 1991.
3. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994 Jan;51(1):8-19. [doi: 10.1001/archpsyc.1994.03950010008002] [Medline: 8279933]
4. Wittchen H, Zhao S, Kessler RC, Eaton WW. DSM-III-R generalized anxiety disorder in the National Comorbidity Survey. Arch Gen Psychiatry 1994 May;51(5):355-364. [doi: 10.1001/archpsyc.1994.03950050015002] [Medline: 8179459]
5. Grigsby AB, Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. Prevalence of anxiety in adults with diabetes: a systematic review. J Psychosom Res 2002 Dec;53(6):1053-1060. [doi: 10.1016/s0022-3999(02)00417-8] [Medline: 12479986]
6. Wilcox G. Insulin and insulin resistance. Clin Biochem Rev 2005 May;26(2):19-39 [FREE Full text] [Medline: 16278749]
7. Gonzalez JS, Peyrot M, McCarl LA, Collins EM, Serpa L, Mimiaga MJ, et al. Depression and diabetes treatment nonadherence: a meta-analysis. Diabetes Care 2008 Dec;31(12):2398-2403 [FREE Full text] [doi: 10.2337/dc08-1341] [Medline: 19033420]
8. American Association of Diabetes Educators. AADE position statement: AADE7™ self-care behaviors. Diabetes Educ 2008;34(3):445-449. [doi: 10.1177/0145721708316625] [Medline: 18535317]
9. Faustino B, Vasco AB, Silva AN, Marques T. Relationships between emotional schemas, mindfulness, self-compassion and unconditional self-acceptance on the regulation of psychological needs. Res Psychother 2020 Sep 07;23(2):442 [FREE Full text] [doi: 10.4081/ripppo.2020.442] [Medline: 33024719]
10. Medina W, Wilson D, de Salvo V, Vannucchi B, de Souza EL, Lucena L, et al. Effects of mindfulness on diabetes mellitus: rationale and overview. Curr Diabetes Rev 2017 Feb 17;13(2):141-147. [doi: 10.2174/1573399812666160607074817] [Medline: 27280721]
11. Tovote KA, Fleer J, Snippe E, Peeters AC, Emmelkamp PM, Sanderman R, et al. Individual mindfulness-based cognitive therapy and cognitive behavior therapy for treating depressive symptoms in patients with diabetes: results of a randomized controlled trial. Diabetes Care 2014 Sep;37(9):2427-2434. [doi: 10.2337/dc13-2918] [Medline: 24898301]
12. Shigaki C, Kruse RL, Mehr D, Sheldon KM, Ge B, Moore C, et al. Motivation and diabetes self-management. Chronic Illn 2010 Sep;6(3):202-214. [doi: 10.1177/1742395310375630] [Medline: 20675362]
13. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
14. Guicciardi M, Lecis R, Anziani C, Corgiolu L, Porru A, Pusceddu M, et al. Type 2 diabetes mellitus, physical activity, exercise self-efficacy, and body satisfaction. An application of the transtheoretical model in older adults. Health Psychol Behav Med 2014 Jan 01;2(1):748-758 [FREE Full text] [doi: 10.1080/21642850.2014.924858] [Medline: 25750816]
15. Menezes MC, Mingoti SA, Cardoso CS, Mendonça RD, Lopes AC. Intervention based on Transtheoretical Model promotes anthropometric and nutritional improvements - a randomized controlled trial. Eat Behav 2015 Apr;17:37-44. [doi: 10.1016/j.eatbeh.2014.12.007] [Medline: 25553558]
16. Alzeidan R, Shata Z, Hassounah MM, Baghdadi LR, Hersi A, Fayed A, et al. Effectiveness of digital health using the transtheoretical model to prevent or delay type 2 diabetes in impaired glucose tolerance patients: protocol for a randomized control trial. BMC Public Health 2019 Nov 21;19(1):1550 [FREE Full text] [doi: 10.1186/s12889-019-7921-8] [Medline: 31752774]
17. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu L, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure (SMS-text adherence support [StAR]): a single-blind, randomized trial. Circulation 2016 Mar 09;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
18. Farmer AJ, Prevost AT, Hardeman W, Craven A, Sutton S, Griffin SJ, SupportAdvice for Medication Trial Group. Protocol for SAMS (Support and Advice for Medication Study): a randomised controlled trial of an intervention to support patients with type 2 diabetes with adherence to medication. BMC Fam Pract 2008 Apr 11;9:20 [FREE Full text] [doi: 10.1186/1471-2296-9-20] [Medline: 18405345]
19. Knickerbocker J, Budd R, Dang B, Chen Q, Colgan E, Hung L, et al. Heterogeneous integration technology demonstrations for future healthcare, IoT, and AI computing solutions. In: Proceedings of the IEEE 68th Electronic Components and Technology Conference (ECTC). 2018 Presented at: IEEE 68th Electronic Components and Technology Conference (ECTC); May 29 -Jun 1, 2018; San Diego, CA, USA. [doi: 10.1109/ectc.2018.00231]
20. Kulyk O, Akker R, Klaassen R, van Gemert-Pijnen L. Personalized virtual coaching for lifestyle support: principles for design and evaluation. Int J Adv Life Sci 2014;6(3 & 4):300-309 [FREE Full text]
21. Kaufman N. Internet and information technology use in treatment of diabetes. Int J Clin Pract Suppl 2010 Mar(166):41-46. [doi: 10.1111/j.1742-1241.2009.02277.x] [Medline: 20377663]
22. O'Brien H. Theoretical perspectives on user engagement. In: Why Engagement Matters. Cham: Springer; 2016.
23. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. Int J Human Comput Stud 2018 Apr;112:28-39. [doi: 10.1016/j.ijhcs.2018.01.004]

24. Michaelides A, Major J, Pienkosz E, Wood M, Kim Y, Toro-Ramos T. Usefulness of a novel mobile diabetes prevention program delivery platform with human coaching: 65-week observational follow-up. JMIR Mhealth Uhealth 2018 May 03;6(5):e93 [FREE Full text] [doi: 10.2196/mhealth.9161] [Medline: 29724709]
25. Baptista S, Wadley G, Bird D, Oldenburg B, Speight J, My Diabetes Coach Research Group. User experiences with a type 2 diabetes coaching app: qualitative study. JMIR Diabetes 2020 Jul 17;5(3):e16692 [FREE Full text] [doi: 10.2196/16692] [Medline: 32706649]
26. Stanger C, Kowatsch T, Xie H, Nahum-Shani I, Lim-Liberty F, Anderson M, et al. A digital health intervention (SweetGoals) for young adults with type 1 diabetes: protocol for a factorial randomized trial. JMIR Res Protoc 2021 Mar 23;10(2):e27109 [FREE Full text] [doi: 10.2196/27109] [Medline: 33620330]
27. Vida Home. Vida. URL: https://www.vida.com/ [accessed 2021-12-29]
28. Tsiouris KM, Tsakanikas VD, Gatsios D, Fotiadis DI. A review of virtual coaching systems in healthcare: closing the loop with real-time feedback. Front Digit Health 2020;2:567502 [FREE Full text] [doi: 10.3389/fdgth.2020.567502] [Medline: 34713040]
29. Sherifali D, Viscardi V, Bai J, Ali RM. Evaluating the effect of a diabetes health coach in individuals with type 2 diabetes. Can J Diabetes 2016 Mar;40(1):84-94. [doi: 10.1016/j.jcjd.2015.10.006] [Medline: 26827684]
30. Tackling type 2 diabetes: try integrated virtual care. Endocrine Web. URL: https://www.endocrineweb.com/news/diabetes/60727-tackling-type-2-diabetes-try-integratedvirtual-care [accessed 2021-12-29]
31. Gender action. International Organization for Standardization. URL: https://www.iso.org/home.html [accessed 2021-12-29]
32. Orsama A, Lähteenmäki J, Harno K, Kulju M, Wintergerst E, Schachner H, et al. Active assistance technology reduces glycosylated hemoglobin and weight in individuals with type 2 diabetes: results of a theory-based randomized trial. Diabetes Technol Ther 2013 Aug;15(8):662-669. [doi: 10.1089/dia.2013.0056] [Medline: 23844570]
33. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. J Med Internet Res 2015 Oct 23;17(10):e240 [FREE Full text] [doi: 10.2196/jmir.4897] [Medline: 26499966]
34. Gabrielli S, Rizzi S, Bassi G, Carbone S, Maimone R, Marchesoni M, et al. Engagement and effectiveness of a healthy-coping intervention via chatbot for university students during the COVID-19 pandemic: mixed methods proof-of-concept study. JMIR Mhealth Uhealth 2021 May 28;9(5):e27965 [FREE Full text] [doi: 10.2196/27965] [Medline: 33950849]
35. Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, et al. Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. BMJ 2013 Mar 26;346:f540 [FREE Full text] [doi: 10.1136/bmj.f540] [Medline: 23444423]
36. Gong E, Baptista S, Russell A, Scuffham P, Riddell M, Speight J, et al. My diabetes coach, a mobile app-based interactive conversational agent to support type 2 diabetes self-management: randomized effectiveness-implementation trial. J Med Internet Res 2020 Nov 05;22(11):e20322 [FREE Full text] [doi: 10.2196/20322] [Medline: 33151154]
37. Rossi MC, Nicolucci A, Lucisano G, Pellegrini F, Di Bartolo P, Miselli V, et al. Impact of the "Diabetes Interactive Diary" telemedicine system on metabolic control, risk of hypoglycemia, and quality of life: a randomized clinical trial in type 1 diabetes. Diabetes Technol Ther 2013 Aug;15(8):670-679. [doi: 10.1089/dia.2013.0021] [Medline: 23844569]
38. Rossi MC, Nicolucci A, Di Bartolo P, Bruttomesso D, Girelli A, Ampudia FJ, et al. Diabetes Interactive Diary: a new telemedicine system enabling flexible diet and insulin therapy while improving quality of life: an open-label, international, multicenter, randomized study. Diabetes Care 2010 Jan;33(1):109-115 [FREE Full text] [doi: 10.2337/dc09-1327] [Medline: 19808926]
39. Bassi G, Mancinelli E, Dell'Arciprete G, Rizzi S, Gabrielli S, Salcuni S. Efficacy of eHealth interventions for adults with diabetes: a systematic review and meta-analysis. Int J Environ Res Public Health 2021 Aug 26;18(17):8982 [FREE Full text] [doi: 10.3390/ijerph18178982] [Medline: 34501572]
40. Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015 Oct;34(10):971-982 [FREE Full text] [doi: 10.1037/hea0000161] [Medline: 25642841]
41. The future of customer experience. RASA. URL: https://rasa.com/ [accessed 2021-12-29]
42. Dinesh P, Sujitha V, Salma C, Srijayapriya B. A review on natural language processing: back to basics. In: Innovative Data Communication Technologies and Application. Singapore: Springer; 2021.
43. Metodo mindfulness. 56 giorni alla felicita. Milan, Italy: Mondadori; 2016.
44. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
45. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
46. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]

47. McGuire BE, Morrison TG, Hermanns N, Skovlund S, Eldrup E, Gagliardino J, et al. Short-form measures of diabetes-related emotional distress: the Problem Areas in Diabetes Scale (PAID)-5 and PAID-1. Diabetologia 2010 Jan;53(1):66-69. [doi: 10.1007/s00125-009-1559-5] [Medline: 19841892]
48. Wellbeing measures in primary health care/the DepCare Project. World Health Organization. Regional Office for Europe. 1998. URL: https://apps.who.int/iris/handle/10665/349766 [accessed 2021-12-29]
49. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
50. O'Brien H, Cairns P. An empirical evaluation of the User Engagement Scale (UES) in online news environments. Inf Process Manag 2015 Jul;51(4):413-427. [doi: 10.1016/j.ipm.2015.03.003]
51. Schrepp M. User Experience Questionnaire Handbook. User Experience Questionnaire. URL: https://www.ueq-online.org/Material/Handbook.pdf [accessed 2022-01-06]
52. The comprehensive R archive network. CRAN. URL: https://cran.r-project.org/ [accessed 2021-12-29]
53. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp; 2016.
54. Karl A, Wisnowski J, Rushing WH. A practical guide to text mining with topic extraction. WIREs Comput Stat 2015 Aug 03;7(5):326-340. [doi: 10.1002/wics.1361]
55. Welbers K, Van Atteveldt W, Benoit K. Text analysis in R. Commun Methods Measures 2017 Nov 02;11(4):245-265. [doi: 10.1080/19312458.2017.1387238]
56. Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, et al. quanteda: an R package for the quantitative analysis of textual data. J Open Source Softw 2018 Oct;3(30):774. [doi: 10.21105/joss.00774]
57. Xiong J, Lipsitz O, Nasri F, Lui LM, Gill H, Phan L, et al. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J Affect Disord 2020 Dec 01;277:55-64 [FREE Full text] [doi: 10.1016/j.jad.2020.08.001] [Medline: 32799105]
58. Healthy coping in diabetes: a guide for program development and implementation. Diabetes Initiative National Program Office. URL: http://diabetesnpo.im.wustl.edu/lessons/documents/HealthyCopinginDiabetesFINAL1-6-10.pdf [accessed 2021-12-29]
59. Von Korff M, Tiemens B. Individualized stepped care of chronic illness. West J Med 2000 Mar;172(2):133-137 [FREE Full text] [doi: 10.1136/ewjm.172.2.133] [Medline: 10693379]

Abbreviations

AADE: American Association of Diabetes Educators
DM: diabetes mellitus
GAD-7: Generalized Anxiety Disorder-7 Items
ML: machine learning
NLU: natural language understanding
ORBIT: Obesity-Related Behavioral Intervention Trials
PAID-5: Problem Areas in Diabetes Scale–Short Form-5 Items
PHQ-9: Patient Health Questionnaire-9 Items
PSS-10: Perceived Stress Scale-10 Items
T1DM: type 1 diabetes mellitus
T2DM: type 2 diabetes mellitus
TTMC: transtheoretical model of change
UE: user engagement
UEQ: User Experience Questionnaire
UES: User Engagement Scale
UX: user experience
VC: virtual coach
WHO-5: The World Health Organization-5 Well-Being Index

Edited by A Kushniruk; submitted 19.07.21; peer-reviewed by D Chrimes, A Joseph; comments to author 10.10.21; revised version received 18.10.21; accepted 07.11.21; published 21.01.22.

Please cite as:
Bassi G, Giuliano C, Perinelli A, Forti S, Gabrielli S, Salcuni S
A Virtual Coach (Motibot) for Supporting Healthy Coping Strategies Among Adults With Diabetes: Proof-of-Concept Study
JMIR Hum Factors 2022;9(1):e32211
URL: https://humanfactors.jmir.org/2022/1/e32211
doi:10.2196/32211
PMID:35060918

Original Paper

Web-Based Tool (FH Family Share) to Increase Uptake of Cascade Testing for Familial Hypercholesterolemia: Development and Evaluation

Hana Bangash[1], MBBS; Ahmed Makkawy[2], BSc; Justin H Gundelach[1], MSc; Alexandra A Miller[1], BSc; Kimberly A Jacobson[1], BA; Iftikhar J Kullo[1], MD

[1]Mayo Clinic, Rochester, MN, United States

[2]Saharafox Creative Agency, Rochester, MN, United States

Corresponding Author:
Iftikhar J Kullo, MD
Mayo Clinic
200 First Street SW
Rochester, MN, 55905
United States
Phone: 1 5072845467
Fax: 1 507 266 9142
Email: kullo.iftikhar@mayo.edu

Abstract

Background: Familial hypercholesterolemia, a prevalent genetic disorder, remains significantly underdiagnosed in the United States. Cascade testing, wherein individuals diagnosed with familial hypercholesterolemia— probands—contact their family members to inform them of their risk for familial hypercholesterolemia, has low uptake in the United States. Digital tools are needed to facilitate communication between familial hypercholesterolemia probands and their family members and to promote sharing of familial hypercholesterolemia–related risk information.

Objective: We aimed to create and evaluate a web-based tool designed to enhance familial communication and promote cascade testing for familial hypercholesterolemia.

Methods: A hybrid type 1 implementation science framework and a user-centered design process were used to develop an interactive web-based tool—FH Family Share—that enables familial hypercholesterolemia probands to communicate information about their familial hypercholesterolemia diagnosis with at-risk relatives. Probands can also use the tool to draw a family pedigree and learn more about familial hypercholesterolemia through education modules and curated knowledge resources. Usability guidelines and standards were taken into account during the design and development of the tool. The initial prototype underwent a cognitive walkthrough, which was followed by usability testing with key stakeholders including genetic counselors and patients with familial hypercholesterolemia. Participants navigated the prototype using the think-aloud technique, and their feedback was used to refine features of the tool.

Results: Key themes that emerged from the cognitive walkthrough were design, format, navigation, terminology, instructions, and learnability. Expert feedback from the cognitive walkthrough resulted in a rebuild of the web-based tool to align it with institutional standards. Usability testing with genetic counselors and patients with familial hypercholesterolemia provided insights on user experience, satisfaction and interface design and highlighted specific modifications that were made to refine the features of FH Family Share. Genetic counselors and patients with familial hypercholesterolemia suggested inclusion of the following features in the web-based tool: (1) a letter-to-family-member email template, (2) education modules, and (3) knowledge resources. Surveys revealed that 6 of 9 (67%) genetic counselors found information within FH Family Share very easy to find, and 5 of 9 (56%) genetic counselors found information very easy to understand; 5 of 9 (56%) patients found information very easy to find within the website, and 7 of 9 (78%) patients found information very easy to understand. All genetic counselors and patients indicated that FH Family Share was a resource worth returning to.

Conclusions: FH Family Share facilitates communication between probands and their relatives. Once informed, at-risk family members have the option to seek testing and treatment for familial hypercholesterolemia.

(JMIR Hum Factors 2022;9(1):e32568) doi:10.2196/32568

KEYWORDS

familial hypercholesterolemia; cascade testing; communication; genetic counselors; digital tools; website; usability; user experience; public health

Introduction

Familial hypercholesterolemia is one of the most common genetic disorders worldwide and is a significant public health burden [1]. With a prevalence of approximately 1 in 250, it is estimated that, in the United States, there are 1.3 million individuals with familial hypercholesterolemia and only 10% have been diagnosed [2,3]. Familial hypercholesterolemia is a treatable disorder, yet due to the lack of awareness, patients remain at significantly increased risk of premature coronary heart disease due to elevated low-density lipoprotein cholesterol levels starting early in life [4-6]. Therefore, increasing familial hypercholesterolemia detection by using cascade testing is important to prevent coronary heart disease and reduce familial hypercholesterolemia–related morbidity and mortality.

Cascade testing, wherein individuals diagnosed with familial hypercholesterolemia (probands) contact their family members and encourage them to get tested for familial hypercholesterolemia, is the most cost-effective method of detecting new cases of familial hypercholesterolemia [7]. Cascade testing has been successfully implemented in a number of countries, most prominently in the Dutch health care system [8]. However, the uptake of cascade testing in the United States is low due to a number of barriers, such as the lack of a centralized and coordinated cascade testing program for familial hypercholesterolemia, the inability of health care providers to directly contact family members due to the Health Insurance Portability and Accountability Act Privacy Rule [9], complex family dynamics, and the burden placed on probands and health care providers in locating and contacting all at-risk family members [10]. In a recent study [11], only 28 of 240 (12%) familial hypercholesterolemia probands were able to enroll a family member for cascade testing, which highlights the low uptake of cascade testing in the United States.

Innovative digital tools have a central role to play in the implementation of genomic medicine by facilitating patient-centered care and decreasing disparities in health care by allowing increased access to care in diverse and underserved communities [12-14]. Well-designed digital tools may also be used to enhance the patient experience by encouraging patient engagement, promoting informed health care–related decisions, and increasing knowledge dissemination [13-15].

To develop digital tools for genomic medicine, it is necessary to obtain patient and provider input, ideally within an implementation science framework, to assess the tool's potential effectiveness as well as institutional and individual level readiness for tool implementation. Patient and provider insights can be used to guide iterative refinements to digital tools and ensure smooth integration into clinical workflows. Hybrid study designs enable elements from both clinical effectiveness research and implementation science research to be blended to serve as a useful framework within which to gather stakeholder feedback and facilitate the translation of digital tools into practice [16].

Methods

Ethics

This study met institutional criteria for Quality Improvement and thus was not subject to review by the Institutional Review Board. The study was conducted from January 2018 to March 2021.

Description of FH Family Share

We used a hybrid type 1 implementation science framework and a collaborative user-centered design process to develop an interactive website intended to facilitate communication between familial hypercholesterolemia probands and their family members. The design was guided by US Department of Health and Human Services Guidelines [17] and International Organization for Standardization Quality Standards for Usability [18]—guidelines that address web design and evaluation as well as user experience optimization for knowledge dissemination. Probands can share their familial hypercholesterolemia diagnosis with relatives via the letter-to-family-member email template. Email allows for faster communication than that using traditional postal mail; the email contains information on the relative likelihood of family members also having familial hypercholesterolemia (given that it is passed on as an autosomal dominant trait) and a recommendation to get tested. The proband can include information about the specific pathogenic variant that was found in their genetic test report, to facilitate family member genetic testing. The website also enables probands to build a family tree utilizing a pedigree tool (AboutMe, Mayo Clinic) to include first-degree relatives and document family members who may be at increased risk for familial hypercholesterolemia. Once a pedigree has been built, it can be accessed each time the user logs on. The web-based tool also has a section, called Learn, with educational modules on familial hypercholesterolemia–related topics: (1) What is Familial Hypercholesterolemia, (2) Familial Hypercholesterolemia Considerations in Children, (3) Genetics, (4) Genetic Testing Frequently Asked Questions, (5) Treatment, and (6) Additional Resources. Through these modules, users (probands or family members) can expand their knowledge on familial hypercholesterolemia and access informative patient education materials and links to other educational tools and websites.

In its current form, FH Family Share will be available on the Mayo Clinic intranet and the internet; all aspects of the web-based tool will be available for public access, with the exception of the pedigree tool which requires the use of a Mayo Clinic patient username and password, to ensure security of protected health information.

Cognitive Walkthrough

The initial proof-of-concept prototype of the web-based tool was built by an external vendor using open-source PHP framework (Yii, version 2.0; Take The Wind). The graphical user interface was designed in HTML5 and CSS3 in conjunction

with Jquery, connected to the MySQL database. It was developed as a responsive design for multiple screen sizes. The initial prototype underwent cognitive walkthrough—a technique applied to evaluate the usability of an app or system in the early stages of design. The cognitive walkthrough was conducted by 3 usability experts to assess exploratory learning—how well an end user can navigate the app for the first time without prior training [19-21]. The usability experts conducted the cognitive walkthrough at the Mayo Clinic Usability Laboratory and participated in 11 tasks that a first-time user was likely to undertake (Multimedia Appendix 1). The session lasted 2 hours and was observed by 5 study team members. Observers gave feedback on any difficulties they noted the experts encounter while navigating the website. The experts and observers discussed the feedback after each task, and points were documented on how to further improve the tool (Multimedia Appendix 2).

Pilot Testing Program With Genetic Counselors

We evaluated user experience, satisfaction, and interface design of FH Family Share with usability testing that was informed by quality standards.[18]. A 1-year pilot testing program was launched in 2 phases (Rochester campus, Mayo Clinic), with genetic counselors as key stakeholders of the web-based tool. Purposive sampling was applied to recruit genetic counselors with varying clinical backgrounds, to obtain information-rich relevant insights on user experience and interface design. Genetic counselors were invited to participate in the usability testing sessions by email. A target sample size of 5 to 8 participants was established based on prior usability studies [22-25], in which approximately 80% to 85% of usability-related concerns were identified from the first 8 participants. We conducted 9 usability testing sessions in phase 1, and we conducted 7 usability testing sessions in phase 2.

Each usability session was 1-hour long; sessions were audiorecorded using a handheld recorder and transcribed using transcription software (version 2018; Otter.ai). Sessions were conducted by a user experience expert (AM) from an external company, to reduce any institutional or workflow driven biases. Study team members (HB and JHG) were also present during each session to observe and take notes. Participating genetic counselors were asked to assess educational content and interface design, as well as to provide insights on how FH Family Share would be integrated with their usual clinical workflows.

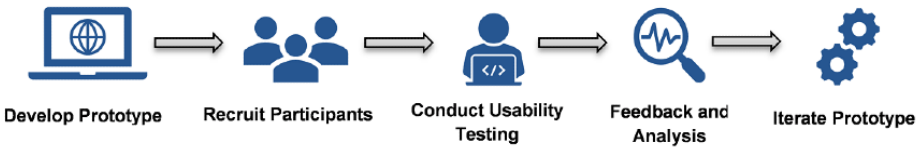
We used a digital platform (version 2011; InVisionApp Inc) to build an interactive clickable prototype of FH Family Share and to conduct user experience testing. The website prototype and 2 case scenarios (Multimedia Appendix 3) were presented to each genetic counselor; each participant was asked to verbalize their thoughts while navigating the prototype [26].

To facilitate feedback, genetic counselors were asked open-ended questions: “How do you feel each of the elements [on the website] might help you in your role?” “What do you think you will find when you click into this page?” “How might you share this website with a patient?” “What are some strengths and weaknesses [of the website]?” “Clicking that print button, what would you imagine would print out?”

Each round of user testing informed iterative refinements and modifications to the features of the prototype. User testing rounds were conducted until saturation was reached (ie, no new feedback was obtained) (Figure 1).

At the end of each usability testing session in phase 1, genetic counselors completed a 7-item satisfaction survey. Based on feedback from genetic counselors in phase 1, the survey was refined, and 4 additional questions were added (Multimedia Appendix 4); therefore, at the end of phase 2, genetic counselors completed an 11-item satisfaction survey.

Figure 1. FH Family Share development and testing work flow.



Evaluation With Patients

To obtain insights from end users, we conducted usability testing sessions over a 3-month period with patients with familial hypercholesterolemia. Patients were recruited using convenience sampling with a target sample size of 8 to 10 participants [22-25]. Patients with a confirmed pathogenic/likely pathogenic genetic variant for familial hypercholesterolemia and who previously participated in familial hypercholesterolemia–related research studies at Mayo Clinic were eligible for inclusion. Patients with a confirmed genetic diagnosis of familial hypercholesterolemia were selected to participate in this study as they had either gone through the process of cascade testing or were currently doing so and could therefore share important feedback and insights on the usefulness and relevance of FH Family Share in facilitating cascade testing. Patients received a request to participate in the study via 1 of 3 methods: the institutional patient portal, a telephone call, or an email. If patients indicated that they were interested in participating, they were contacted by phone to schedule the usability testing session.

Due to the COVID-19 pandemic, usability testing sessions were conducted using videoconferencing software (Zoom Video Communications Inc). Each session lasted 1 hour. Both audio and video from sessions were recorded; audiorecordings were transcribed using Office Word (version 2021; Microsoft Inc). Each session was conducted by a user experience expert (AM); study team members (HB, AA, and JHG) were also present during each session. Patients were emailed a link to access the clickable prototype of the web-based tool and were asked to

think aloud when navigating the prototype. Patients did not receive any compensation for participating in the study.

At the start of each session, patients were prompted to describe their individual diagnostic journeys for familial hypercholesterolemia and provide insights on their family structure: “When were you diagnosed with FH (provide your age at the time of diagnosis or the calendar year)?” “Tell me about your experience receiving your FH diagnosis (mode and duration of communication, formats it was received in, and any patient education material you may have been given at the time).” “How did the way in which you received your diagnosis make you feel?” “What resources, if any, did you use to look up more information about FH?” “How did you find these resources (did you reach out to providers, search the internet etc)?” “How many first-degree relatives (parents, siblings, children) do you have?” “Did you share any of the information pertaining to your diagnosis with family members (if so, in what context or format)?” “If you did not share your diagnosis with family members, were there any barriers that prevented you from doing so?” “How many of your family members, that you know of, have completed testing for FH?” At the end of each session, patients were asked to complete an 11-item survey using web-based software (QualtricsXM) (Multimedia Appendix 5).

Results

Cognitive Walkthrough

Key themes that emerged from the cognitive walkthrough were design, format, navigation, terminology, instructions and learnability (Multimedia Appendix 6).

Pilot Testing Program With Genetic Counselors

Usability Testing

Of 13 genetic counselors who were contacted, 9 consented to participate in phase 1. In phase 2, all 7 genetic counselors who were contacted agreed to participate (Multimedia Appendix 7).

Usability testing with genetic counselors resulted in key workflow insights and highlighted specific modifications that could be made to FH Family Share. Genetic counselors found FH Family Share to be a welcome tool for use in disseminating knowledge and information to patients and facilitating communication on health risks information between patients and their family members. In current practice, once patients receives a diagnosis of familial hypercholesterolemia, they must use postal mail to send templated letters to family members informing them of their increased risk and recommending screening for familial hypercholesterolemia; genetic counselors indicated that the template email would aid patients in sharing their diagnosis and other pertinent information with family members and health care providers. In phase 1, 7 of 9 (78%) genetic counselors highlighted the need for a template email in which patients or genetic counselors could enter the name of the gene, pathogenic variant, and laboratory that conducted the genetic testing. Genetic counselors indicated that such information pertaining to the pathogenic variant would be valuable in enabling health care providers of family members determine which genetic testing was needed. Based on this feedback, the FH Family Share prototype was modified to include entry fields in the template email (Figure 2) that could be updated with relevant information on the pathogenic variant obtained from the patient’s genetic test report (this health information is not stored on the website).

Genetic counselors highlighted the importance of enabling patients to learn more about familial hypercholesterolemia because this would increase the likelihood of patients recommending cascade testing to their family members. However, all genetic counselors suggested rearranging the order of the content into a sequence that would be more in line with how they were likely to approach the conversation during their usual clinical workflows (Figure 3).

Additionally, 7 of 9 (78%) genetic counselors who participated in phase 1 of the pilot program recommended replacing stock photography and detailed medical images with patient-friendly material. This feedback led the study team to collaborate with graphic designers to develop medical illustrations for the physical manifestations of familial hypercholesterolemia, such as corneal arcus and tendon xanthomas [27].

Genetic counselors appreciated the integration of the pedigree tool with FH Family Share, particularly for its value in enabling patients to build their own pedigree and identify family members who may need to be screened for familial hypercholesterolemia, thereby facilitating earlier detection and treatment. Genetic counselors also valued the additional resources included on the website, such as links to patient education materials and videos.

Figure 2. Screenshot of the template email on FH Family Share.

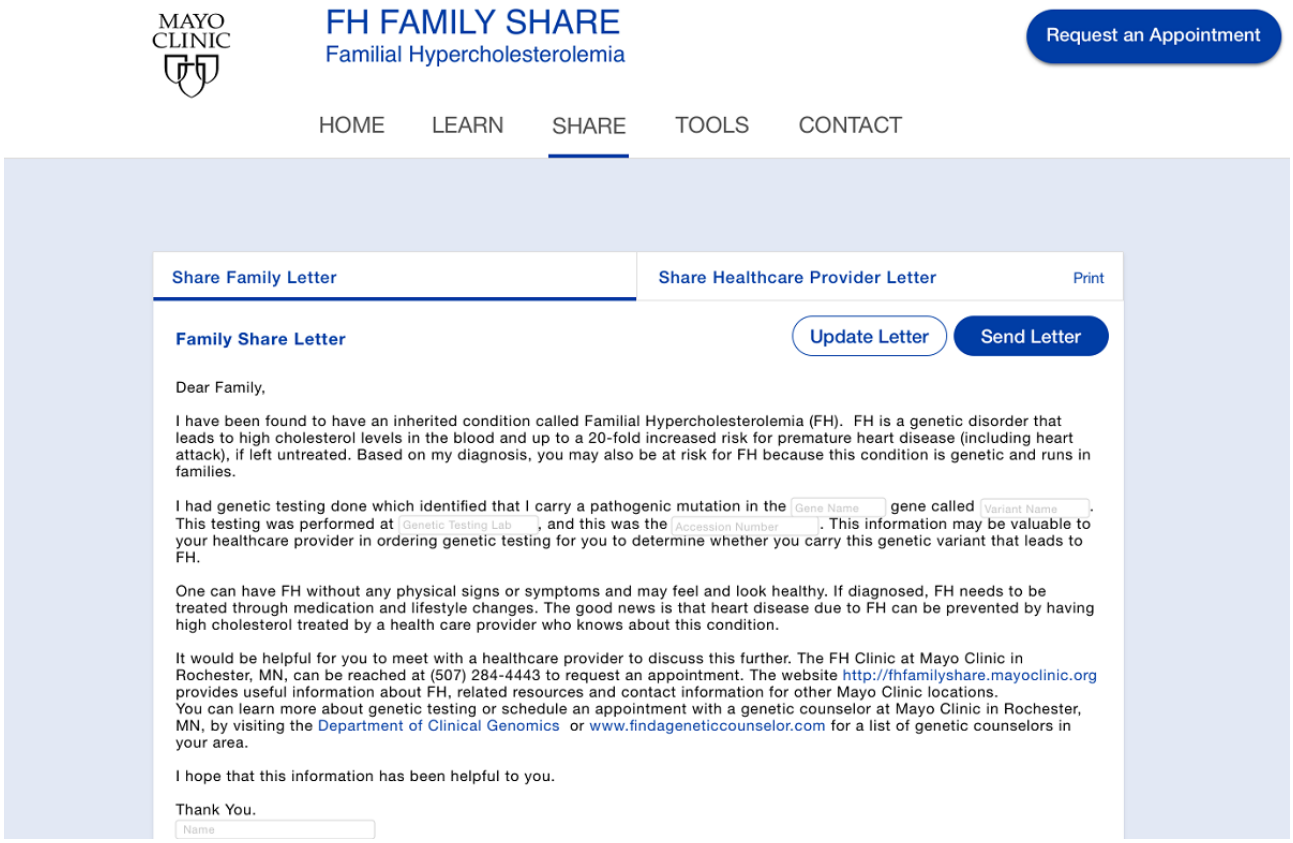


Figure 3. Sequence of educational topics.

Survey Results

The majority of the genetic counselors found information within the FH Family Share website very easy to find (phase 1: 6/9, 67%; phase 2: 6/7, 86%) and very easy to understand (phase 1: 5/9, 56%; phase 2: 6/7, 86%). In both phases of the pilot program (Table 1 and Table 2), all genetic counselors responded that the FH Family Share website was a resource worth returning to. In response to the questions added to the survey in phase 2, 5 of 7 (71%) genetic counselors agreed or completely agreed that the FH Family Share tool would ease their workflow during patient encounters, all genetic counselors agreed or completely agreed that the tool would likely improve follow-up patient care, 4 of 7 (57%) genetic counselors indicated that patients would likely use the Learn modules, and 2 of 7 (29%) genetic counselors indicated that patients were likely to send a letter to family members communicating their familial hypercholesterolemia diagnosis.

Table 1. Satisfaction survey responses from genetic counselors.

Item	Phase 1 (n=9), n (%)	Phase 2 (n=7), n (%)
Overall, the information that you were asked to assess within the FH Family Share website was		
Very easy to find	6 (67)	6 (86)
Somewhat easy to find	3 (33)	1 (14)
Overall, the information that you found within the FH Family Share website was		
Very easy to understand	5 (56)	6 (86)
Somewhat easy to understand	4 (44)	1 (14)
Is the FH Family Share website a resource worth returning to?		
Yes	9 (100)	7 (100)
No	0 (0)	0 (0)
The FH Family Share website will ease my workflow in a patient encounter[a]		
Completely agree	—[b]	2 (29)
Agree	—	3 (43)
Neither agree nor disagree	—	2 (29)
The FH Family Share website is likely to improve follow-up patient care[a]		
Completely agree	—	3 (43)
Agree	—	4 (57)
Neither agree nor disagree	—	0 (0)
As a provider I feel the patient is most likely to[a,c]		
Use Learn modules as a knowledge resource	—	4 (57)
Build a family tree using AboutMe	—	0 (0)
Calculate risk of heart attack	—	1 (14)
Send a letter to family members	—	2 (29)
Use website information to discuss familial hypercholesterolemia with family members	—	0 (0)
Do you find the website figures/images/diagrams useful?		
Yes	—	6 (86)
No	—	0 (0)
Other[d]	—	1 (14)

[a]The question was added for phase 2.

[b]No data for phase 1.

[c]Only 1 option could be selected.

[d]"Some were useful; more scientific diagrams were not as helpful from a patient perspective" [Participant 1, genetic counselor].

Table 2. Representative comments from genetic counselors in response to free-text questions. (Participant 4 did not provide free-text responses in phase 2.)

Item	Representative comments	
	Phase 1	Phase 2
What did you like most about the FH Family Share website?	"Solid basic information. Good patient reading level." [Participant 4] "The education materials, getting contact info for the clinic." [Participant 5] "The importance of informing family members and assistance to do this, especially the email." [Participant 8]	"It was interactive and user-friendly. I think this will provide patients with a good tool to learn more about FH and how to easily share this information with their relatives." [Participant 2] "Nice resource for families including the ability to share the resource and family letter within the website" [Participant 6]
What did you like least about the FH Family Share website?	"Ensure appropriate nomenclature is being used for heterozygous vs. homozygous FH." [Participant 1] "Add more infographics about inheritance, genetics etc" [Participant 3] "The order of the modules under learn and the ambiguousness of 'Discover'" [Participant 8]	"Images could use work" [Participant 1] "There is a lot of clicking into different tabs" [Participant 3]
What additional information or functionality would you like to see on the FH Family Share website?	"Personalization of email tool." [Participant 2] "Next steps" for family members, how to find provider, how to share info with family." [Participant 3]	"Use for other health care providers w/ limited experience in FH, incorporate Dutch Lipid Network, etc" [Participant 7]
What more could we do to improve the FH Family Share website?	"More meaningful Images." [Participant 6] "A little more information about the About Me tool would be helpful prior to them starting it." [Participant 7]	"Again, nothing further to add. Looks great!" [Participant 2] "Make thing as concise as possible" [Participant 3] "Define terms or link to where they are defined." [Participant 5]

FH Family Share Evaluation With Patients

Usability Testing

Of 28 patients who were contacted, 13 responded with interest, and of those, 9 patients consented and participated in the usability testing sessions (Multimedia Appendix 8). The target sample size for participant recruitment was achieved. Patients who participated in the usability testing sessions had between 4 and 11 first-degree relatives; the majority of the patients (5 of 9, 56%) had shared their familial hypercholesterolemia diagnosis with all first-degree relatives, and additionally, 8 of the 9 patients had at least 1 or more first-degree relatives known to have undergone cholesterol testing or genetic testing.

Feedback obtained from patient usability testing sessions covered both content and interface design. Patients highlighted the need for a new section to be included in the *Learn* section of FH Family Share that would address the implications of familial hypercholesterolemia for children of variant-positive parents, the age at which children should be screened and treated, as well as the implications of having familial hypercholesterolemia when planning to start a family. Patients wanted to know the next best steps in these situations and emphasized the need to highlight the emotional motivation for getting tested for familial hypercholesterolemia and sharing a diagnosis with family members; children and grandchildren were identified as being the greatest motivators for patients. Additional feedback was obtained on the need to expand upon the information provided on proprotein convertase subtilisin/kexin type 9 inhibitors in the Treatment section by including details on how to obtain insurance approvals. Patients also shared input on interface design including need for a larger clearer font, to make content easier to read (using bullets instead of paragraphs), to use a less formal tone in the family letter, to add biographies to health care provider names on the Contact page, and to include more images and illustrations to make the website more interesting and engaging. Patient feedback led to further iterations of FH Family Share to incorporate their input.

Survey Results

Analysis of the satisfaction surveys revealed that 5 of 9 (56%) patients found information within the FH Family Share website very easy to find, and 7 of 9 (78%) found information very easy to understand (Table 3). When asked to share their perspective on how the website would impact patient care, 2 of 9 (22%) patients responded that it would significantly improve care, while 5 of 9 (56%) responded that it would somewhat improve care. The majority of patients agreed that the website would make it easier for them to understand and share their familial hypercholesterolemia diagnosis (completely agree: 5/9, 56%; agree: 4/9, 44%). When asked about the one activity they would most likely perform while using the website, 3 of 9 (33%) patients indicated that they would use the risk calculator to determine risk of heart attack. All patients agreed that the FH Family Share website was a resource worth returning to.

Patients also gave responses to 4 free-text survey questions (Table 4).

Table 3. Satisfaction survey responses from patients.

Survey questions	Patient responses, n (%)
Overall, the information that you were asked to assess within the FH Family Share website was (n=9)	
Very easy to find	5 (56)
Somewhat easy to find	4 (44)
Overall, the information that you found within the FH Family Share website was (n=9)	
Very easy to understand	7 (78)
Somewhat easy to understand	2 (22)
How will the FH Family Share website impact patient care/follow-up care? (n=8)[a]	
Significantly improve	2 (22)
Somewhat improve	5 (56)
Neither improve nor worsen	1 (11)
The FH Family Share website will make it easier for patients like myself to understand and share an FH diagnosis (n=9)	
Completely Agree	5 (56)
Agree	4 (44)
As a patient, I would be most likely to (n=9)	
Use the Learn modules as a knowledge resource	2 (22)
Build a family tree using AboutMe	0 (0)
Calculate risk of a heart attack	3 (33)
Send a letter to family members	2 (22)
Use website information to discuss FH with family members	2 (22)
Do you find the FH Family Share website figures/images/diagrams useful? (n=9)	
Yes	8 (89)
No	1 (11)
Is the FH Family Share website a resource worth returning to? (n=9)	
Yes	9 (100)
No	0 (0)

[a]Participant 6 did not provide a response to this survey question.

Table 4. Representative comments from patients with familial hypercholesterolemia in response to free-text questions.

Survey question	Representative comments
What did you like most about the FH Family Share website?	"The design intent is solid, and I did find more information regarding genetic variants than I had found doing research from other sources (Mayo, NIH, CDC)" [Participant 2] "It's nice to have all this information clearly compiled to be able to return and reference, especially for someone like me who has dealt with this diagnosis for most of my life but continue to learn more about it and how it will impact my health and potentially that of my family as I get older." [Participant 5]
What did you like least about the FH Family Share website?	"There was no content that I disliked - the flow of the site could be improved a bit, as reviewed in the session." [Participant 1] "Wish more of the links would have been active - but overall, the view and ease to move around was solid." [Participant 2] "It was boring and not very visually stimulating." [Participant 8]
What additional information or functionality would you like to see on the FH Family Share website?	"The risk stratification tool based on medications would be an awesome addition to this website." [Participant 1] "As mentioned, perhaps more information geared toward younger patients, starting a family, thinking about getting children tested/when/why, and how risk can continue to be minimized over a lifetime." [Participant 5] "Maybe interactive images." [Participant 9]
What more could we do to improve the FH Family Share website?	"Create more emotional connection to why treatment is so important - perhaps by emphasizing risks. So many young people think they are invincible - as I did, as well - somehow make it more real, since you don't feel the effects of FH until it may be too late." [Participant 2] "Pictures, diagrams, maybe address the holistic challenges folks with FH have in common." [Participant 7]

Discussion

Principal Findings

Given the ubiquitous use of electronic health records, the internet, and smartphones, digital tools can serve a central role in the delivery and implementation of genomic medicine [10,14,28]. Familial hypercholesterolemia is underdiagnosed, and there is limited uptake of cascade testing in the United States [3,29,30]. FH Family Share was designed to facilitate communication between familial hypercholesterolemia probands and their relatives, and to increase the uptake of cascade testing. It was iteratively refined through usability testing with genetic counselors and patients with familial hypercholesterolemia. The target sample size for participant recruitment was achieved in both phases of the pilot testing program, and based on study team consensus, additional participants were not recruited as feedback saturation was reached and majority of the usability issues had been identified.

Current guidelines [31,32] for familial hypercholesterolemia screening recommend the identification of a proband and cascade testing of family members, starting with first-degree relatives. Cascade testing for familial hypercholesterolemia has been identified as the most cost-effective strategy for identifying new familial hypercholesterolemia cases [33,34]. In traditional cascade testing, a proband sends a templated letter, using postal mail, to inform relatives of their diagnosis and to encourage them to get tested; however, there is no way to determine whether relatives receive the letter and seek testing for familial hypercholesterolemia. This approach to cascade testing has low uptake in the United States, in part due to the inability of health care providers to directly contact family members and the burden placed on probands to undertake this task, and in part due to the lack of an established national program with cascade testing that is centralized, coordinated, and aligned with local and regional needs and resources [35,36]. Although countries with nationalized health care systems have been able to implement centralized cascade testing programs, the United States is limited by several barriers. Well-designed usable digital tools can be used to bridge the gap in communication and increase uptake of cascade testing.

Usability testing is an integral part of the field of human–computer interaction and is used to evaluate interactive health apps to ensure effective design and development strategies, with a particular focus on the concept of user-centered design [37-41]. Genetic counselors and patients with familial hypercholesterolemia are the intended end users of the web-based tool, and feedback from these groups informed iterative refinements to FH Family Share after each phase of testing. Our extensive usability testing with genetic counselors revealed FH Family Share to be a resource worth returning to, with easy to find and easy to understand content that would likely improve patient follow-up care. Genetic counselors indicated that the tool was something they could share on their computer screens with patients during pre- and post-genetic testing counseling sessions. Genetic counselors also highlighted that the template email would assist patients in sharing their familial hypercholesterolemia diagnosis with relatives but that it needed to include sections in which either they (the genetic counselors) or the patients could insert gene and variant names, as this would be useful to family members when scheduling testing. Orlando et al [42,43] employed a similar process to develop a family health history and decision support tool, in which usability testing conducted with 10 genetic counselors resulted in a number of adaptations to the final tool, such as changes to interface design, navigation, and content, prior to its deployment.

Usability feedback from patients with familial hypercholesterolemia revealed that the majority of patients found content easy to understand, agreed that the tool could help them share their familial hypercholesterolemia diagnosis with family members, and was a resource worth returning to. A novel insight gained from patient usability testing was the

importance of understanding the motivation for patients seeking genetic testing for familial hypercholesterolemia. Patients revealed that concern for their children, grandchildren, and family planning were the primary motivators. They highlighted that FH Family Share should contain an education module focused on considerations for children and family planning. This feedback led to the addition of a new topic module. Similar feedback was obtained in a study [44] conducted to identify motivators and barriers to cascade testing in families with familial hypercholesterolemia. Study participants indicated that they informed relatives of their risk of familial hypercholesterolemia to protect them from heart disease and allow them to make appropriate lifestyle changes [44].

FH Family Share will be implemented at all Mayo Clinic sites and the Mayo Clinic Health System and will also be available for public access on the internet. Once deployed, the tool's content and resources will be updated as new knowledge emerges. The impact and metrics of FH Family Share will be assessed using a pilot implementation study. The findings described herein could be used to create similar digital apps for other genomic disorders.

Limitations

Usability testing sessions were limited to genetic counselors and patients with familial hypercholesterolemia—input from family members of patients with familial hypercholesterolemia would be useful in ensuring effective implementation of the tool in nonclinical settings. Additionally, FH Family Share was not integrated with the electronic health record to enable ease of access in remote or underserved communities which may have different electronic health record systems in place. In some instances, feedback obtained from genetic counselors and patients with familial hypercholesterolemia varied on the same topic, for instance, most patients indicated the need for more illustrations on FH Family Share to increase engagement and interest, while genetic counselors suggested limiting the use of illustrations unless they served a specific purpose; feedback from both groups was harmonized by adding curated illustrations to increase patient engagement. Due to the COVID-19 pandemic, usability testing sessions conducted with patients were virtual and nonverbal cues could not be observed; however, screen sharing along with open-ended questions allowed us to obtain sufficient feedback which was often consistent across participants. The sample sizes for genetic counselors and patients with familial hypercholesterolemia were modest; however, the numbers were in line with those of previous usability studies [22-25] and allowed for thematic saturation to be reached.

Conclusions

FH Family Share can be used by probands to communicate their familial hypercholesterolemia diagnosis with family members, encouraging them to seek testing for familial hypercholesterolemia. Such a tool has the potential to increase uptake of cascade testing thereby allowing for earlier detection and treatment of familial hypercholesterolemia.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Task scenarios used for the cognitive walkthrough conducted by usability experts.
[DOCX File , 20 KB - humanfactors_v9i1e32568_app1.docx]

Multimedia Appendix 2
Design and development specifications of FH Family Share.
[DOCX File , 13 KB - humanfactors_v9i1e32568_app2.docx]

Multimedia Appendix 3
Case scenarios.
[DOCX File , 13 KB - humanfactors_v9i1e32568_app3.docx]

Multimedia Appendix 4
Satisfaction survey (genetic counselors).
[DOCX File , 16 KB - humanfactors_v9i1e32568_app4.docx]

Multimedia Appendix 5
Survey (patients).
[DOCX File , 16 KB - humanfactors_v9i1e32568_app5.docx]

Multimedia Appendix 6
Themes and representative quotations.
[DOCX File , 14 KB - humanfactors_v9i1e32568_app6.docx]

Multimedia Appendix 7
Demographic characteristics (genetic counselors).
[DOCX File , 14 KB - humanfactors_v9i1e32568_app7.docx]

Multimedia Appendix 8
Demographic characteristics (patients).
[DOCX File , 14 KB - humanfactors_v9i1e32568_app8.docx]

References

1. Safarova MS, Kullo IJ. Lessening the burden of familial hypercholesterolemia using health information technology. Circ Res 2018 Jan 05;122(1):26-27 [FREE Full text] [doi: 10.1161/CIRCRESAHA.117.312319] [Medline: 29301842]
2. Nordestgaard BG, Chapman MJ, Humphries SE, Ginsberg HN, Masana L, Descamps OS, European Atherosclerosis Society Consensus Panel. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. Eur Heart J 2013 Dec;34(45):3478-390a [FREE Full text] [doi: 10.1093/eurheartj/eht273] [Medline: 23956253]
3. Knowles JW, Rader DJ, Khoury MJ. Cascade screening for familial hypercholesterolemia and the use of genetic testing. JAMA 2017 Jul 25;318(4):381-382 [FREE Full text] [doi: 10.1001/jama.2017.8543] [Medline: 28742895]
4. Safarova MS, Liu H, Kullo IJ. Rapid identification of familial hypercholesterolemia from electronic health records: the SEARCH study. J Clin Lipidol 2016;10(5):1230-1239. [doi: 10.1016/j.jacl.2016.08.001] [Medline: 27678441]
5. Lee C, Rivera-Valerio M, Bangash H, Prokop L, Kullo IJ. New case detection by cascade testing in familial hypercholesterolemia: a systematic review of the literature. Circ Genom Precis Med 2019 Nov;12(11):e002723. [doi: 10.1161/CIRCGEN.119.002723] [Medline: 31638829]
6. Bangash H, Khan F, He B, Arce M, Kullo IJ. Use of Twitter to promote awareness of familial hypercholesterolemia. Circ Genom Precis Med 2019 Jul;12(7):e002550 [FREE Full text] [doi: 10.1161/CIRCGEN.119.002550] [Medline: 31306057]
7. Knowles JW, Rader DJ, Khoury MJ. Cascade screening for familial hypercholesterolemia and the use of genetic testing. JAMA 2017 Jul 25;318(4):381-382 [FREE Full text] [doi: 10.1001/jama.2017.8543] [Medline: 28742895]
8. Umans-Eckenhausen MA, Defesche JC, Sijbrands EJ, Scheerder RL, Kastelein JJ. Review of first 5 years of screening for familial hypercholesterolaemia in the Netherlands. Lancet 2001 Jan 20;357(9251):165-168. [doi: 10.1016/S0140-6736(00)03587-X] [Medline: 11213091]
9. The HIPAA Privacy Rule. US Department of Health and Human Services. URL: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html [accessed 2022-02-03]
10. Kullo IJ. Familial hypercholesterolemia: a reportable disorder. Circulation 2020 Nov 24;142(21):1999-2001 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.120.050548] [Medline: 33226870]
11. Ajufo E, deGoma EM, Raper A, Yu KD, Cuchel M, Rader DJ. A randomized controlled trial of genetic testing and cascade screening in familial hypercholesterolemia. Genet Med 2021 May 26:1697-1704. [doi: 10.1038/s41436-021-01192-z] [Medline: 34040191]
12. Bombard Y, Hayeems RZ. How digital tools can advance quality and equity in genomic medicine. Nat Rev Genet 2020 Sep;21(9):505-506 [FREE Full text] [doi: 10.1038/s41576-020-0260-x] [Medline: 32601319]
13. Bangash H, Sutton J, Gundelach JH, Pencille L, Makkawy A, Elsekaily O, et al. Deploying clinical decision support for familial hypercholesterolemia. ACI Open 2020 Dec 31;04(02):e157-e161. [doi: 10.1055/s-0040-1721489]
14. Bangash H, Pencille L, Gundelach JH, Makkawy A, Sutton J, Makkawy L, et al. An implementation science framework to develop a clinical decision support tool for familial hypercholesterolemia. J Pers Med 2020 Jul 23;10(3):67. [doi: 10.3390/jpm10030067] [Medline: 32717811]
15. Shickh S, Rafferty SA, Clausen M, Kodida R, Mighton C, Panchal S, Incidental Genomics Study Team. The role of digital tools in the delivery of genomic medicine: enhancing patient-centered care. Genet Med 2021 Jun;23(6):1086-1094. [doi: 10.1038/s41436-021-01112-1] [Medline: 33654192]
16. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
17. U.S. Department of Health and Human Services. 2006. URL: https://www.usability.gov/sites/default/files/documents/guidelines_book.pdf [accessed 2022-02-03]
18. ISO 9241-11 Usability: definitions and concepts. Part 11. International Organization for Standardization. 2018. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed 2021-06-26]
19. Rieman J. An automated cognitive walkthrough. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1991 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27-May 2; New Orleans, Louisiana p. 427-428. [doi: 10.1145/108844.108986]

20. Rieman J, Franzke M, Redmiles D. Usability evaluation with the cognitive walkthrough. In: Conference Companion on Human Factors in Computing Systems. 1995 Presented at: Human Factors in Computing Systems Conference; May 7-11; Denver, Colorado p. 387-388.
21. John BE, Packer H. Learning and using the cognitive walkthrough method: a case study approach. 1995 Presented at: Human Factors in Computing Systems Conference; 1995; Denver, Colorado p. 429-436. [doi: 10.1145/223904.223962]
22. Georgsson M, Staggers N. Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. J Am Med Inform Assoc 2016 Jan;23(1):5-11 [FREE Full text] [doi: 10.1093/jamia/ocv099] [Medline: 26377990]
23. Macefield R. How to specify the participant group size for usability studies: a practitioners guide. J Usability Stud 2009;5:45.
24. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 Feb;37(1):56-76. [doi: 10.1016/j.jbi.2004.01.003] [Medline: 15016386]
25. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough. Hum Factors 2016 Nov 23;34(4):457-468. [doi: 10.1177/001872089203400407]
26. Kastner M, Lottridge D, Marquez C, Newton D, Straus SE. Usability evaluation of a clinical decision support tool for osteoporosis disease management. Implement Sci 2010 Dec 10;5:96 [FREE Full text] [doi: 10.1186/1748-5908-5-96] [Medline: 21143978]
27. Safarova MS, Liu H, Kullo IJ. Rapid identification of familial hypercholesterolemia from electronic health records: The SEARCH study. J Clin Lipidol 2016;10(5):1230-1239. [doi: 10.1016/j.jacl.2016.08.001] [Medline: 27678441]
28. McCarthy JJ, McLeod HL, Ginsburg GS. Genomic medicine: a decade of successes, challenges, and opportunities. Sci Transl Med 2013 Jun 12;5(189):189sr4. [doi: 10.1126/scitranslmed.3005785] [Medline: 23761042]
29. Knowles JW, O'Brien EC, Greendale K, Wilemon K, Genest J, Sperling LS, et al. Reducing the burden of disease and death from familial hypercholesterolemia: a call to action. Am Heart J 2014 Dec;168(6):807-811 [FREE Full text] [doi: 10.1016/j.ahj.2014.09.001] [Medline: 25458642]
30. Kullo IJ, Bailey KR. Design of a controlled trial of cascade screening for hypercholesterolemia: the (CASH) study. J Pers Med 2018 Aug 23;8(3):1-10 [FREE Full text] [doi: 10.3390/jpm8030027] [Medline: 30142941]
31. Goldberg AC, Hopkins PN, Toth PP, Ballantyne CM, Rader DJ, Robinson JG, et al. Familial hypercholesterolemia: screening, diagnosis and management of pediatric and adult patients: clinical guidance from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. J Clin Lipidol 2011;5(3):133-140. [doi: 10.1016/j.jacl.2011.03.001] [Medline: 21600517]
32. Migliara G, Baccolini V, Rosso A, D'Andrea E, Massimi A, Villari P, et al. Familial hypercholesterolemia: a systematic review of guidelines on genetic testing and patient management. Front Public Health 2017;5:252 [FREE Full text] [doi: 10.3389/fpubh.2017.00252] [Medline: 28993804]
33. Marks D, Wonderling D, Thorogood M, Lambert H, Humphries SE, Neil HAW. Cost effectiveness analysis of different approaches of screening for familial hypercholesterolaemia. BMJ 2002 Jun 01;324(7349):1303 [FREE Full text] [doi: 10.1136/bmj.324.7349.1303] [Medline: 12039822]
34. Goldman L, Goldman PA, Williams LW, Weinstein MC. Cost-effectiveness considerations in the treatment of heterozygous familial hypercholesterolemia with medications. Am J Cardiol 1993 Sep 30;72(10):75D-79D. [doi: 10.1016/0002-9149(93)90015-5] [Medline: 8213502]
35. HHS National Conferences on the HIPAA Privacy Rule. US Department of Health and Human Services. 2021 Dec 07. URL: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html [accessed 2020-06-25]
36. Bell DA, Pang J, Burrows S, Bates TR, van Bockxmeer FM, Hooper AJ, et al. Effectiveness of genetic cascade screening for familial hypercholesterolaemia using a centrally co-ordinated clinical service: an Australian experience. Atherosclerosis 2015 Mar;239(1):93-100. [doi: 10.1016/j.atherosclerosis.2014.12.036] [Medline: 25585028]
37. 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]
38. Earthy J, Jones BS. ISO standards for user-centered design and the specification of usability. In: Usability in Government Systems. Amsterdam: Elsevier; 2012:267-283.
39. Brunner J, Chuang E, Goldzweig C, Cain CL, Sugar C, Yano EM. User-centered design to improve clinical decision support in primary care. Int J Med Inform 2017 Dec;104:56-64 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.05.004] [Medline: 28599817]
40. Martinez W, Threatt AL, Rosenbloom ST, Wallston KA, Hickson GB, Elasy TA. A patient-facing diabetes dashboard embedded in a patient web portal: design sprint and usability testing. JMIR Hum Factors 2018 Sep 24;5(3):e26 [FREE Full text] [doi: 10.2196/humanfactors.9569] [Medline: 30249579]
41. Khairat S, Marc D, Crosby W, Al Sanousi A. Reasons for physicians not adopting clinical decision support systems: critical analysis. JMIR Med Inform 2018 Apr 18;6(2):e24 [FREE Full text] [doi: 10.2196/medinform.8912] [Medline: 29669706]
42. Wu RR, Orlando LA, Himmel TL, Buchanan AH, Powell KP, Hauser ER, et al. Patient and primary care provider experience using a family health history collection, risk stratification, and clinical decision support tool: a type 2 hybrid controlled

implementation-effectiveness trial. BMC Fam Pract 2013 Aug 06;14:111 [FREE Full text] [doi: 10.1186/1471-2296-14-111] [Medline: 23915256]
43. Orlando LA, Buchanan AH, Hahn SE, Christianson CA, Powell KP, Skinner CS, et al. Development and validation of a primary care-based family health history and decision support program (MeTree). N C Med J 2013 Jul 01;74(4):287-296. [doi: 10.18043/ncm.74.4.287]
44. Wurtmann E, Steinberger J, Veach PM, Khan M, Zierhut H. Risk communication in families of children with familial hypercholesterolemia: identifying motivators and barriers to cascade screening to improve diagnosis at a single medical center. J Genet Couns 2018 Dec 14;28(1):50-58. [doi: 10.1007/s10897-018-0290-0]

Edited by A Kushniruk; submitted 04.08.21; peer-reviewed by D Rader, D Chrimes; comments to author 15.10.21; revised version received 29.10.21; accepted 07.11.21; published 15.02.22.

Please cite as:
Bangash H, Makkawy A, Gundelach JH, Miller AA, Jacobson KA, Kullo IJ
Web-Based Tool (FH Family Share) to Increase Uptake of Cascade Testing for Familial Hypercholesterolemia: Development and Evaluation
JMIR Hum Factors 2022;9(1):e32568
URL: https://humanfactors.jmir.org/2022/1/e32568
doi:10.2196/32568
PMID:35166678

Original Paper

The Effects of Introducing a Mobile App–Based Procedural Logbook on Trainee Compliance to a Central Venous Catheter Insertion Accreditation Program: Before-and-After Study

Robert Tamblyn[1], MBBS; Jorge Brieva[1,2], MBBS; Madeleine Cain[1], MBBS; F Eduardo Martinez[1,2], MD

[1]Division of Critical Care, John Hunter Hospital, Hunter New England Local Health District, Newcastle, Australia

[2]School of Medicine and Public Health, Faculty of Health and Medicine, University of Newcastle, Newcastle, Australia

Corresponding Author:
Robert Tamblyn, MBBS
Division of Critical Care
John Hunter Hospital
Hunter New England Local Health District
Lookout Rd, New Lambton Heights
Newcastle, 2305
Australia
Phone: 61 432032151
Email: robert.d.tamblyn@gmail.com

Abstract

Background: To reduce complications associated with central venous catheter (CVC) insertions, local accreditation programs using a supervised procedural logbook are essential. To increase compliance with such a logbook, a mobile app could provide the ideal platform for training doctors in an adult intensive care unit (ICU).

Objective: The aim of this paper was to compare trainee compliance with the completion of a logbook as part of a CVC insertion accreditation program, before and after the introduction of an app-based logbook.

Methods: This is a retrospective observational study of logbook data, before and after the introduction of a purpose-built, app-based, electronic logbook to complement an existing paper-based logbook. Carried out over a 2-year period in the adult ICU of the John Hunter Hospital, Newcastle, NSW, Australia, the participants were ICU trainee medical officers completing a CVC insertion accreditation program. The primary outcome was the proportion of all CVC insertions documented in the patients' electronic medical records appearing as logbook entries. To assess logbook entry quality, we measured and compared the proportion of logbook entries that were approved by a supervisor and contained a supervisor's signature for the before and after periods. We also analyzed trainee participation before and after the intervention by comparing the total number of active logbook users, and the proportion of first-time users who logged 3 or more CVC insertions.

Results: Of the 2987 CVC insertions documented in the electronic medical records between April 7, 2019, and April 6, 2021, 2161 (72%) were included and separated into cohorts before and after the app's introduction. Following the introduction of the app-based logbook, the percentage of CVC insertions appearing as logbook entries increased from 3.6% (38/1059) to 20.5% (226/1102; $P<.001$). There was no difference in the proportion of supervisor-approved entries containing a supervisor's signature before and after the introduction of the app, with 76.3% (29/38) and 83.2% (188/226), respectively ($P=.31$). After the introduction of the app, there was an increase in the percentage of active logbook users from 15.3% (13/85) to 62.8% (54/86; $P<.001$). Adherence to one's logbook was similar in both groups with 60% (6/10) of first-time users in the before group and 79.5% (31/39) in the after group going on to log at least 3 or more CVCs during their time working in ICU.

Conclusions: The addition of an electronic app-based logbook to a preexisting paper-based logbook was associated with a higher rate of logbook compliance in trainee doctors undertaking an accreditation program for CVC insertion in an adult ICU. There was a large increase in logbook use observed without a reduction in the quality of logbook entries. The overall trainee participation also improved with an observed increase in active logbook users and no reduction in the average number of entries per user following the introduction of the app. Further studies on app-based logbooks for ICU procedural accreditation programs are warranted.

(JMIR Hum Factors 2022;9(1):e35199) doi:10.2196/35199

KEYWORDS

logbook; education; training; central venous catheter; CVC; intensive care; smartphone; mobile phone; mobile apps; mHealth; mobile health; accreditation program; digital health; digital record

Introduction

Background

Central venous catheter (CVC) insertion is a common procedure performed in the intensive care unit (ICU) with up to 78% of ICU patients requiring insertion of a CVC during their admission [1]. CVC insertion complications occur frequently and can be serious. Mechanical complications include arterial injury, hematoma, and pneumothorax, while delayed complications can include thrombosis and catheter-related bloodstream infections [2]. Mechanical complications have been reported in 2.1% of all subclavian and 1.4% of all internal jugular CVC insertions [3].

In teaching hospitals, trainees perform a large portion of this aspect of care. In a 2017 UK multicenter audit, 62% of all CVC insertions were performed by trainees [4]. This is important to note because complications are more common in less experienced hands [5]. Training and accreditation programs help to reduce the rate of both mechanical complications and catheter-related bloodstream infections associated with less experienced operators [6,7]. While theoretical knowledge can be attained via educational modules, junior doctors need adequate experience and exposure to achieve procedural competency [8].

Logbooks are used to facilitate training and supervision across many fields of postgraduate medicine. In Australia and New Zealand, logbooks are a mandatory requirement of the Royal Australasian College of Surgeons, the Royal Australasian College of Physicians, and the Australasian College for Emergency Medicine [9-11]. Some evidence-based guidelines, such as those published in the British Journal of Anaesthesia, strongly recommend the use of logbooks in the training of CVC insertion and maintenance [12]. Logbooks help clarify learning objectives and can facilitate communication between the trainee and their clinical supervisor. They encourage procedural supervision of trainees and help supervisors know when a trainee doctor is competent at performing a procedure independently [13].

Paper-based logbooks have historically had poor compliance as well as incomplete and invalid data entry [14]. On the other hand, electronic logbooks have shown benefits including more efficient data access and an increased ability for supervisors to monitor trainees [15]. In many parts of the world, electronic logbooks have replaced paper-based logbooks, with desktop and website applications being used as their user interface [16,17]. Despite its advantages, a computer-based platform can have inconveniences of its own, with immediate accessibility not always available to the logbook user and their supervisor. In recent years, smartphones have become ubiquitous in hospitals, with medicine-related mobile phone apps being used daily for education and training and as a clinical aid [18]. It is possible that a smartphone app could provide a convenient platform for an electronic logbook for ICU-based procedures such as CVC insertion.

Objectives

This study aims to determine if in a tertiary-level ICU, the introduction of an app-based logbook, when compared to a paper-based logbook, was associated with an increase in trainee compliance with the logbook component of a CVC accreditation program.

Methods

Design

This is a single-center, before-and-after, retrospective observational study of prospective data of logbook compliance.

Setting and Participants

This study was conducted in the John Hunter Hospital ICU between April 7, 2019, and April 6, 2021. The John Hunter Hospital ICU is a 27-bed mixed medical and surgical unit in a tertiary, university-affiliated teaching hospital. It has over 2000 admissions per year and has an average of approximately 1100 CVC insertions annually.

The John Hunter Hospital is accredited for ICU general, as well as cardiothoracic surgery, neurosurgery, and trauma ICU training by the College of Intensive Care Medicine of Australia and New Zealand. Its junior medical roster accommodates 34 junior doctors at different levels of training at any one time, who are completing 3-month, 6-month, or 12-month rotations as part of their postgraduate training.

The ICU established a formal accreditation program for CVC insertion in 2013, introduced to improve patient safety and reduce complication rates. In 2017, the program incorporated a mandatory web-based training module followed by a practical component.

To be accredited as an independent operator for a specific type of procedure, trainees need to be supervised completing a prescribed number of CVC insertions. Each insertion must be logged by trainees in a communal, paper-based logbook. Each entry must contain a minimum level of information and be signed by a valid supervisor. A major limitation of the paper-based logbook has been its poor accessibility to both trainees and supervisors, contributing to poor trainee compliance.

Intervention

In April 2020, a smart phone app was developed using AppSheet, a low code mobile app builder [19]. The app was introduced to the John Hunter Hospital ICU to increase the use of the procedural logbook, previously limited to a paper-based logbook kept in a communal folder. The app combined a read-only daily unit roster with an interactive, updatable procedural logbook.

Interface design focused on rapid data entry, minimum text requirements, and field-level automation. Individuals would enter de-identified patient data regarding the CVC insertion procedure using their mobile device, and the data would be immediately uploaded to a secure, central, web-based database. The app could be installed on both iPhone and Android operating systems. The app was made available to all doctors working in the unit from April 6, 2020, following a brief introductory session. The session was repeated each 3-month term for all new trainees rotating through the unit. The paper-based logbook remained available to all trainees, and they could use either the paper-based or app-based logbooks to create new entries.

Variables

The primary outcome was compliance with the CVC insertion logbook, as indicated by the number of logbook entries made by trainees as a proportion of the total number of CVC insertions documented in the patients' electronic medical record (EMR) before and after the introduction of the app-based logbook.

The secondary outcomes were as follows: (1) to assess the quality of the information entered in the logbook by measuring the number of supervisor-approved logbook entries, evidenced by having a supervisor's signature; (2) to assess trainee participation by comparing the number of active logbook book users in the before and after periods; and (3) to assess trainee adherence by measuring the proportion of first-time logbook users who logged 3 or more entries.

Data Collection

Data were collected from the EMR for all standard CVCs, hemodialysis catheters, and peripherally inserted central catheters inserted in the unit from April 7, 2019, until April 6, 2021. The data collected included the following: type of line inserted, insertion time, ICU admission time, patient's medical record number, and insertion site.

Our inclusion criteria were the following: (1) CVCs (standard central lines, hemodialysis catheters, and peripherally inserted central catheter lines); (2) inserted within the ICU; (3) in adult patients.

Our exclusion criteria were the following: (1) invasive lines other than a CVC (eg, arterial line, pulmonary artery catheter, extracorporeal membrane oxygenation catheter, and rapid infusion lines); (2) CVCs inserted in patients younger than 18 years old; and (3) any CVC inserted outside of the ICU (eg, emergency department, operating theatres, wards, and prehospital).

Each CVC insertion documented in the EMR that met the inclusion criteria and had none of the exclusion criteria was recorded. The paper-based and app-based logbooks were then interrogated looking for the relevant information about the CVC insertion each logbook entry was meant to correspond with. The data points of the EMR and logbook entries that had to match were as follows: (1) patient's medical record number, (2) line type, (3) insertion site, and (4) date of insertion. Only if all these variables were the same for both the EMR and the logbook was the CVC insertion defined as having a corresponding logbook entry.

For each insertion having a corresponding logbook entry, we recorded additional data obtained from the logbook, including the proceduralist's name, the presence of a supervisor's signature, and in which logbook, paper, or app did the entry appeared. If the same entry appeared in both paper and digital logbooks, its source was recorded as paper. Using hospital records, we determined variables about the proceduralist including their level of training and the amount of time working at John Hunter Hospital ICU. A first-time logbook user was defined as any trainee making a logbook entry in their first term working in the John Hunter Hospital ICU.

All data, including those of the patients' and the trainees' who performed the procedure, were de-identified and entered in a purpose-built, password-protected, electronic database. Study numbers were assigned to patients and trainees to minimize the risk of de-identification.

We analyzed the database to determine the number of CVC insertions in EMR having a corresponding logbook entry in the before and after period, the presence of a supervisor's signature, and the details of the trainee who performed the insertion.

Study Size

For study size, 365 days before and after the introduction of the app were included to allow the equal representation of all parts of the training year.

Data Analysis

Descriptive statistics are reported as fractions and percentages. Chi-square and Fisher exact tests were used to compare categorical variables. A 2-sided *P* value of <.05 was considered statistically significant.

Ethics

The Hunter New England Human Research Ethics Committee approved an exemption for ethical review for the study given its negligible risk (authorization AU202108-07).

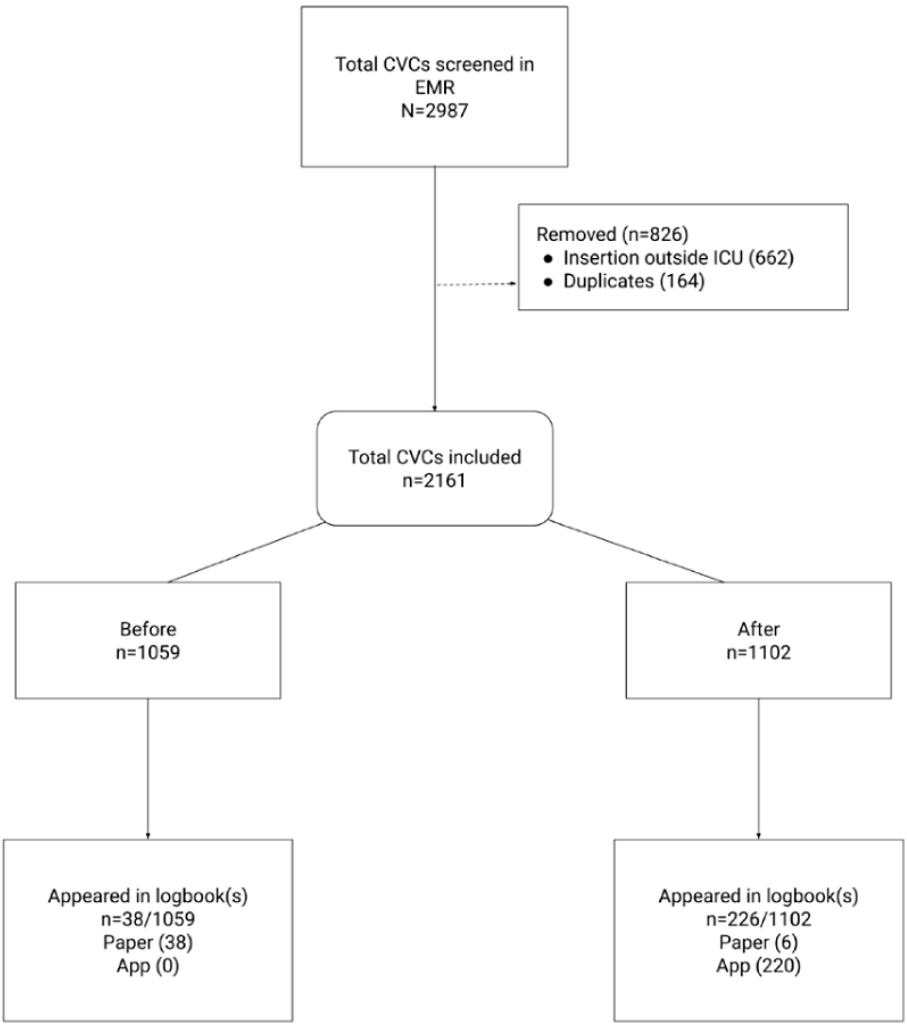
Permission to access the paper logbook was granted by the director of the unit as part of a quality assurance. The trainees provided implied consent when installing the app for all data entered to be collected and analyzed.

Results

Data Collection

A total of 2987 CVC insertions recorded in the EMR between April 7, 2019, and April 6, 2021, were screened. Of these, 2161 were included, with 1059 in the before group and 1102 in the after group. There was a total of 270 insertions that had a corresponding logbook entry, 44 in the paper logbook and 226 in the app-based logbook (Figure 1).

Figure 1. CONSORT diagram. CVC: central venous catheter; EMR: electronic medical record; ICU: intensive care unit.



The types of lines inserted and their insertion locations were well matched at baseline in the before and after groups (Table 1).

Table 1. Central lines recorded in the electronic medical record in patients admitted to the intensive care unit before and after the introduction of the app-based logbook.

Line type and insertion site	Before	After
Total number of lines n	1059	1102
Standard CVC,[a] n (%)		
External jugular vein	2 (0.2)	6 (0.5)
Femoral	108 (10.2)	168 (15.2)
Internal jugular vein	377 (35.6)	394 (35.8)
Other	0 (0)	3 (0.3)
Subclavian	169 (16.0)	120 (10.9)
PICC,[b] n (%)		
Basilic	222 (21.0)	233 (21.1)
Brachial	54 (5.1)	37 (3.4)
Cephalic	38 (3.6)	34 (3.1)
Hemodialysis catheter, n (%)		
Femoral	27 (2.5)	26 (2.4)
Internal jugular vein	54 (5.1)	80 (7.3)
Other	2 (0.2)	2 (0.2)
Subclavian	14 (1.3)	5 (0.5)

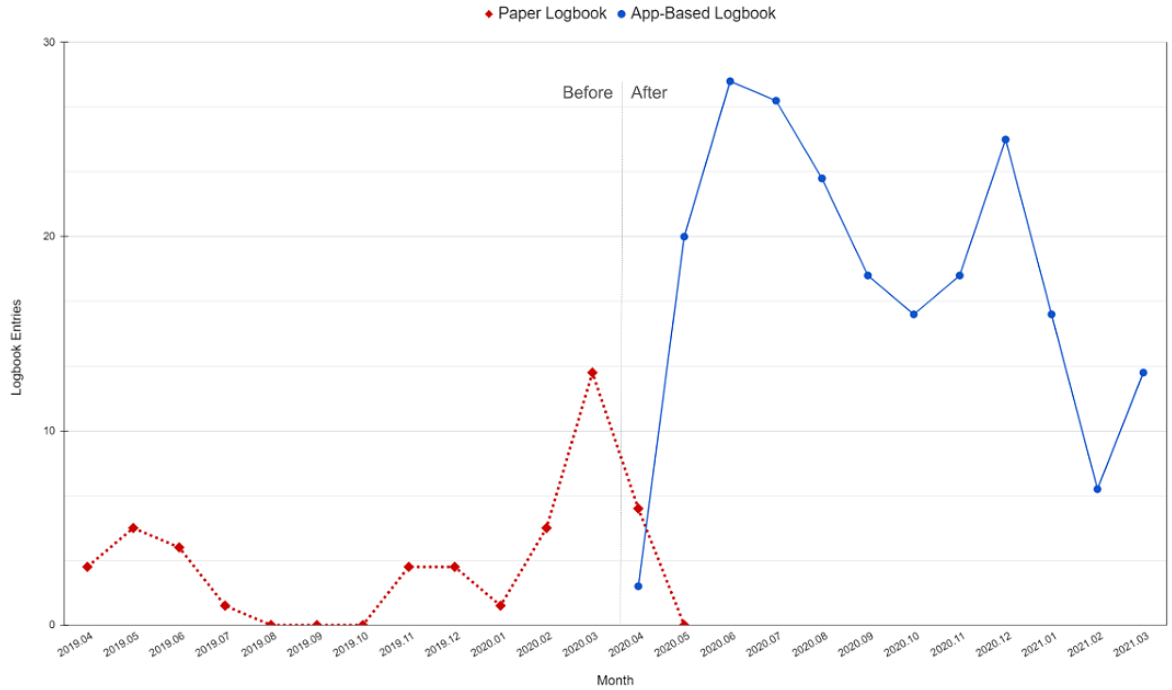
[a]CVC: central venous catheter.

[b]PICC: peripherally inserted central catheter.

Study Outcomes

The rate of logbook entries increased after the introduction of the app-based logbook (Figure 2).

Figure 2. Combined paper-based and app-based central venous catheter logbook entries per month.



There was a statistically significant increase in the percentage of logbook entries out of the total number of CVC insertions documented in EMR, from 3.6% (38/1059) in the before group to 20.5% (226/1102) in the after group (P<.001; Table 2).

Table 2. Percentage of central venous catheter insertions with corresponding logbook entries before and after the introduction of the app-based logbook.

	Before	After	*P* value
Total number of CVCs,[a] n	1059	1102	N/A[b]
Corresponding logbook entry, n (%)	38 (3.6)	226 (20.5)	<.001
No corresponding logbook entry, n (%)	1021 (96.4)	876 (79.5)	<.001

[a]CVC: central venous catheter.

[b]N/A: not applicable.

There was no difference in the proportion of entries containing a supervisor's signature from the total number of logbook entries before and after the introduction of the app (P=.31; odds ratio 0.65, 95% CI 0.29-1.49; Table 3).

Table 3. Percentage of logbook entries before and after the introduction of the app-based logbook containing a valid signature.

	Before	After	Odds ratio	95% CI	*P* value
Total number of logbook entries, n	38	226	N/A[a]	N/A	N/A
Entries with valid supervisor signature, n (%)	29 (76.3)	188 (83.2)	0.65	0.29-1.49	.31

[a]N/A: not applicable.

The number of unique trainees was 146, with 85 counted in the "before" group and 86 in the "after" group. Moreover, 25 trainees were counted in both groups having worked in the John Hunter Hospital ICU during both periods. The training level of the participating trainees in both groups is summarized in Table 4.

Table 4. John Hunter Hospital intensive care unit junior doctors by level of training in the before and after time periods.

Trainees	Before, n/N (%)	After, n/N (%)
Senior registrar	4/15 (26)	0/12 (0)
Registrar	14/25 (56)	15/28 (53)
SRMO[a]	39/41 (95)	34/43 (79)
RMO[b]	4/4 (100)	9/10 (90)

[a]SRMO: senior resident medical officer.

[b]RMO: resident medical officer.

There was an overall increase in the rate of active logbook users, from 15.3% (13/85) to 62.8% (54/86) of all trainees. Of these active users, 10 (76.9%) qualified as first-time users in the "before" group and 39 (72.2%) in the "after" group. There was no significant change in adherence between groups with 60% (6/10) first-time users in the "before" group and 79.5% (31/39) in the "after" group going on to log at least 3 or more CVCs during their time working in the unit (P=.20; odds ratio 0.39, 95% CI 0.09-1.71; Table 5)

Senior resident medical officers had the greatest increase in the rate of logbook adoption following the introduction of the app-based logbook, going from 5.9% (5/85) to 37.2% (32/86) of all trainees in their first term making at least 1 logbook entry (absolute risk increase 31.3%).

Table 5. Adoption and adherence to a logbook. Total number of logbook users, first-time logbook users, and first-time users completing 3 or more entries at John Hunter Hospital intensive care unit.

Level of training	Before	After	Odds ratio	95% CI	*P* value
Unique trainees, n	85	86	N/A[a]	N/A	N/A
Active users, n (%)					
Total	13 (15.3)	54 (62.8)	0.11	0.05-0.22	<.001
Senior registrar	3 (3.5)	4 (4.7)			
Registrar	4 (4.7)	10 (11.6)			
SRMO[b]	5 (5.9)	32 (37.2)			
RMO[c]	1 (1.2)	8 (9.3)			
First-time users (3 or more entries) by total first-time users, n/N (%)					
Total	6/10 (60)	31/39 (79.5)	0.39	0.09-1.71	.20
Senior registrar	1/2 (50)	0/0 (0)			
Registrar	0/2 (0)	6/7 (85.7)			
SRMO	4/5 (80)	21/24 (87.5)			
RMO	1/1 (100)	4/8 (50)			

[a]N/A: not applicable.

[b]SRMO: senior resident medical officer.

[c]RMO: resident medical officer.

Discussion

Principal Findings

In this retrospective before-and-after study, the addition of a mobile app-based logbook to an existing paper logbook resulted in increased compliance among trainees to the logbook component of a CVC insertion accreditation program without reducing the quality and completeness of logbook entries. This increase occurred without there being an increase in the total number of CVC insertions documented in EMR. There was also an increase in the number of active logbook users while maintaining the same level of adherence.

There are several potential explanations for this observed increase in logbook compliance among our study participants. Smartphones have become ubiquitous in hospitals, with many clinicians using mobile apps as a resource for instant access to information and as clinical decision-making tools [18,20].

We have seen earlier studies show widespread uptake of electronic procedural logbooks following their introduction among trainee doctors [15,21,22]. The addition of a smartphone app makes the digital logbook immediately available for rapid data entry, saving the clinician additional time and energy and allowing them to comply with the completion of their logbook.

There was no difference in the proportion of logbook entries with a valid supervisor signature in the before and after periods, demonstrating that the quality and completeness of the logbook entry as perceived by the supervisors was not reduced, nor was the level of overall supervisor interaction with the logbook process. This means that the addition of a mobile app-based logbook increased the quantity of entries while maintaining quality and supervisor participation.

Junior trainee uptake was higher following the introduction of the app, with the total number of active logbook users during the “before” and “after” periods significantly increasing. This indicates that the increased rate of logbook entries was due to more trainees using the logbook rather than the same number of trainees using the logbook more frequently. It reflects a higher rate of adoption among trainees, which may be due to the convenience and availability of the app, but also due to a generational aspect with most trainees in the study likely to have grown up with digital devices as part of daily life [23]. The increased uptake was more apparent in the senior resident medical officer groups who tended to consist of more junior doctors, and less so in the slightly more senior registrar groups, potentially less conducive to change and acceptance of a new process.

While the rate of new users can give some indication of the app’s performance, it can be confounded by a high “churn” rate; the rate of subscribers, in this case to a logbook, who discontinue use early after their first interaction [24]. We observed an increase in the number of first-time logbook users without compromising the number going on to log multiple entries. Features in the app such as recognition of procedural milestones and automated emails may have contributed to a low churn rate among trainees subscribing to the app-based logbook.

How Findings Are Different From Previous Knowledge

There are some important differences between previous research and our study. The outcome measures of other studies did not compare relative numbers of CVC insertions logged before and after the introduction of an electronic-based system. Outcome measures in previous studies included absolute numbers of logbook entries, quality of data recorded, and survey-based trainee feedback. Many were conducted using surgical trainees

among whom procedural logbooks had been a long established training requirement [15,21]. Although 1 study reviewed logbooks already available as mobile phone apps [18], all others focused on website-based logbooks. Moreover, some studies were conducted before smartphones were widely in use [15,16,21].

New and Relevant Data the Study Provides

This is the first before-and-after study directly comparing trainee compliance to a procedural logbook in an ICU setting. Although the benefits of electronic logbooks have been studied previously, this is the first time a mobile app has been observed as an intervention.

Meaning and Implications of the Study

While we have known about the utility of logbooks to facilitate practical aspects of medical training for decades, their application in real world clinical environments has been limited. The inconvenience and time-consuming aspects of accessing and maintaining a physical record is thought to be responsible for the low compliance rate of trainees to their logbooks. This is especially true in a busy ICU environment, where carrying extra documentation or accessing a desktop computer may not be feasible.

This study has shown a significant increase in compliance among trainees in an ICU working toward accreditation in CVC insertion when given the opportunity to use an app compared with a paper logbook, while maintaining the same level of supervision.

Although this was a single-center study, these findings could be generalized to other similar units to ours because the app would be the same everywhere.

Limitations

As a before-and-after study, the difference between the cohorts may be due to factors other than just the introduction of the logbook app. The introductory sessions for the app may have increased awareness and enthusiasm toward the use of logbooks. There may also be a component of "gizmo idolatry" among junior staff; a conviction that a more technological solution to a problem is intrinsically better than a less technological one [25].

Additionally, the study only measured surrogate outcomes. Patient-centered outcomes, namely CVC related complications, were not directly tested. Trainee and supervisor satisfaction were not addressed.

Unanswered Questions

A longer-term study is required to reduce the short-term effects of early adoption of a new technology. A multiple unit study is required to test the external validity of this study and for the scalability of such an app-based logbook. Future research is needed to assess an association between logbook compliance and a reduction in CVC-related complications.

Conclusions

In this before-and-after study, when compared with a previous paper-based system, the addition of a mobile app-based logbook improved the compliance of junior doctors undertaking a CVC insertion accreditation program in an adult intensive care unit, at a tertiary level teaching hospital. There was a marked increase in the number of logbook entries observed as a proportion of all CVC insertions in the unit. Despite this increase, the quality and completeness of entries was not affected. We also observed an increase in the total number of active logbook users after the introduction of the app, with no effect on the average number of logbook entries per user. The more junior trainees accounted for most of this observed increase, potentially signaling a greater willingness to adopt new technologies than their more advanced colleagues. Further studies on app-based logbooks in critical settings are warranted to examine patient outcomes and trainee satisfaction. With digital tools becoming increasingly common in teaching hospitals, it is important to continue to research their impact on medical training and education.

Conflicts of Interest

RT was involved in the development process of the logbook app; however, there are no financial competing interests. The app is implemented on the AppSheet platform, is not monetized, and remains available to all medical staff at the John Hunter Hospital intensive care unit.

References

1. Vincent J, Bihari DJ, Suter PM, Bruining HA, White J, Nicolas-Chanoin M, et al. The Prevalence of Nosocomial Infection in Intensive Care Units in Europe. JAMA 1995 Aug 23;274(8):639. [doi: 10.1001/jama.1995.03530080055041]
2. McGee DC, Gould MK. Preventing Complications of Central Venous Catheterization. N Engl J Med 2003 Mar 20;348(12):1123-1133. [doi: 10.1056/nejmra011883]
3. Parienti J, Mongardon N, Mégarbane B, Mira J, Kalfon P, Gros A, 3SITES Study Group. Intravascular Complications of Central Venous Catheterization by Insertion Site. N Engl J Med 2015 Sep 24;373(13):1220-1229. [doi: 10.1056/NEJMoa1500964] [Medline: 26398070]
4. Wong AV, Arora N, Olusanya O, Sharif B, Lundin RM, Dhadda A, First Intensive Care National Audit Project (ICNAP-1) group. Insertion rates and complications of central lines in the UK population: A pilot study. J Intensive Care Soc 2018 Feb;19(1):19-25 [FREE Full text] [doi: 10.1177/1751143717722914] [Medline: 29456597]
5. Kusminsky RE. Complications of central venous catheterization. J Am Coll Surg 2007 Apr;204(4):681-696. [doi: 10.1016/j.jamcollsurg.2007.01.039] [Medline: 17382229]

6. Warren DK, Zack JE, Mayfield JL, Chen A, Prentice D, Fraser VJ, et al. The effect of an education program on the incidence of central venous catheter-associated bloodstream infection in a medical ICU. Chest 2004 Nov;126(5):1612-1618. [doi: 10.1378/chest.126.5.1612] [Medline: 15539735]
7. Cherry RA, West CE, Hamilton MC, Rafferty CM, Hollenbeak CS, Caputo GM. Reduction of central venous catheter associated blood stream infections following implementation of a resident oversight and credentialing policy. Patient Saf Surg 2011 Jun 03;5(1):15 [FREE Full text] [doi: 10.1186/1754-9493-5-15] [Medline: 21639916]
8. Bannister SL, Hilliard RI, Regehr G, Lingard L. Technical skills in paediatrics: a qualitative study of acquisition, attitudes and assumptions in the neonatal intensive care unit. Med Educ 2003 Dec;37(12):1082-1090. [doi: 10.1111/j.1365-2923.2003.01711.x] [Medline: 14984114]
9. Watters DAK, Green AJ, van Rij A. Requirements for trainee logbooks. ANZ J Surg 2006 Mar;76(3):181-184. [doi: 10.1111/j.1445-2197.2006.03663.x] [Medline: 16626362]
10. Assessments and Requirements - FACEM Trainees Enrolled Prior to 2022. Australian College for Emergency Medicine. URL: https://acem.org.au/Content-Sources/Training/How-the-FACEM-Training-Program-works/FACEM-Trainees-Enrolled-Prior-to-2022/Assessments-and-Requirements-FACEM-Trainees-Enro [accessed 2021-11-04]
11. New Basic Training Program. The Royal Australasian College of Physicians. URL: https://www.racp.edu.au/trainees/basic-training/curricula-renewal/new-basic-training-program [accessed 2021-11-04]
12. Moureau N, Lamperti M, Kelly LJ, Dawson R, Elbarbary M, van Boxtel AJH, et al. Evidence-based consensus on the insertion of central venous access devices: definition of minimal requirements for training. Br J Anaesth 2013 Mar;110(3):347-356 [FREE Full text] [doi: 10.1093/bja/aes499] [Medline: 23361124]
13. Schüttpelz-Brauns K, Narciss E, Schneyinck C, Böhme K, Brüstle P, Mau-Holzmann U, et al. Twelve tips for successfully implementing logbooks in clinical training. Med Teach 2016 Jun;38(6):564-569 [FREE Full text] [doi: 10.3109/0142159X.2015.1132830] [Medline: 26841068]
14. Hosseini M, Irajpour A, Hayrabedian A, Khusravi M, Rarani S, Ghadami A. A study of the views of bachelor students in operating room on evaluation of clinical performance using paper logbook, Isfahan University of Medical Sciences. Ann Trop Med Public Health 2017;10(6):1668. [doi: 10.4103/atmph.atmph_576_17]
15. Aphinives P. Implementation of electronic logbook for trainees of general surgery in Thailand. J Med Assoc Thai 2013 Jan;96(1):47-51. [Medline: 23720977]
16. Achuthan R, Grover K, MacFie J. A critical evaluation of the electronic surgical logbook. BMC Med Educ 2006 Mar 01;6(1):15 [FREE Full text] [doi: 10.1186/1472-6920-6-15] [Medline: 16509987]
17. Cheng A, Saba T, Duncan T, Grundy S, Evison M. Pulmonary passport: a service evaluation study of a standardised web-based procedure logbook to aid specialist respiratory training and appraisal. BMJ Open Respir Res 2020 Nov;7(1):e000690. [doi: 10.1136/bmjresp-2020-000690] [Medline: 33184058]
18. Adam A, Spencer K, Moon S, Jacub I. Smartphone, Smart Surgeon, what about a 'Smart Logbook'? S Afr J Surg 2016 Jun;54(2):36-42. [Medline: 28240502]
19. Make an App with No-Code. AppSheet. URL: https://www.appsheet.com [accessed 2021-11-01]
20. Wiechmann W, Kwan D, Bokarius A, Toohey SL. There's an App for That? Highlighting the Difficulty in Finding Clinically Relevant Smartphone Applications. West J Emerg Med 2016 Mar;17(2):191-194 [FREE Full text] [doi: 10.5811/westjem.2015.12.28781] [Medline: 26973750]
21. Brouwer R, Kiroff G. Computer-based logbook for surgical registrars. ANZ J Surg 2002 Jan 14;72(1):57-61. [doi: 10.1046/j.1445-2197.2002.02296.x] [Medline: 11906426]
22. Ahmed M, Saba T, Duncan T, Grundy S, Evison M. The Pulmonary Passport: Uptake of a regional electronic procedure logbook in Respiratory Medicine Internet. European Respiratory Journal 2018;52(Suppl 62):1645. [doi: 10.1183/13993003.congress-2018.OA1645]
23. Lerchenfeldt S, Attardi SM, Pratt RL, Sawarynski KE, Taylor TAH. Twelve tips for interfacing with the new generation of medical students: iGen. Med Teach 2021 Nov;43(11):1249-1254. [doi: 10.1080/0142159X.2020.1845305] [Medline: 33174808]
24. Kwon H, Kim HH, An J, Lee J, Park YR. Lifelog Data-Based Prediction Model of Digital Health Care App Customer Churn: Retrospective Observational Study. J Med Internet Res 2021 Jan 06;23(1):e22184 [FREE Full text] [doi: 10.2196/22184] [Medline: 33404511]
25. Leff B, Finucane TE. Gizmo idolatry. JAMA 2008 Apr 16;299(15):1830-1832. [doi: 10.1001/jama.299.15.1830] [Medline: 18413879]

Abbreviations

CVC: central venous catheter
EMR: electronic medical record
ICU: intensive care unit

Edited by A Kushniruk; submitted 24.11.21; peer-reviewed by A Montilha; comments to author 08.01.22; revised version received 12.01.22; accepted 20.01.22; published 07.03.22.

Please cite as:
Tamblyn R, Brieva J, Cain M, Martinez FE
The Effects of Introducing a Mobile App–Based Procedural Logbook on Trainee Compliance to a Central Venous Catheter Insertion Accreditation Program: Before-and-After Study
JMIR Hum Factors 2022;9(1):e35199
URL: https://humanfactors.jmir.org/2022/1/e35199
doi:10.2196/35199
PMID:35051900

Original Paper

Understanding Cardiology Practitioners' Interpretations of Electrocardiograms: An Eye-Tracking Study

Mohammed Tahri Sqalli[1], MSc; Dena Al-Thani[1], PhD; Mohamed B Elshazly[2,3], MD; Mohammed Al-Hijji[4,5], MD; Alaa Alahmadi[6], PhD; Yahya Sqalli Houssaini[7], MD

[1]Information and Computing Technology Division, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar

[2]Department of Cardiac Electrophysiology, Cleveland Clinic, Cleveland, OH, United States

[3]Johns Hopkins Ciccarone Center, Heart and Vascular Institute, Johns Hopkins Medicine, Baltimore, MD, United States

[4]Interventional & Structural Cardiology Division, Heart Hospital, Hamad Medical Corporation, Doha, Qatar

[5]Weill Cornell Medical College, Cornell University, Doha, Qatar

[6]College of Computer Science and Engineering, Taibah University, Yanbu, Saudi Arabia

[7]Department of Medicine, Faculty of Medecine and Pharmacy, Mohammed V University, Rabat, Morocco

Corresponding Author:
Mohammed Tahri Sqalli, MSc
Information and Computing Technology Division
College of Science and Engineering
Hamad Bin Khalifa University
Qatar Foundation
Doha, 34110
Qatar
Phone: 974 50588170
Email: mtahrisqalli@hbku.edu.qa

Abstract

Background: Visual expertise refers to advanced visual skills demonstrated when performing domain-specific visual tasks. Prior research has emphasized the fact that medical experts rely on such perceptual pattern-recognition skills when interpreting medical images, particularly in the field of electrocardiogram (ECG) interpretation. Analyzing and modeling cardiology practitioners' visual behavior across different levels of expertise in the health care sector is crucial. Namely, understanding such acquirable visual skills may help train less experienced clinicians to interpret ECGs accurately.

Objective: This study aims to quantify and analyze through the use of eye-tracking technology differences in the visual behavior and methodological practices for different expertise levels of cardiology practitioners such as medical students, cardiology nurses, technicians, fellows, and consultants when interpreting several types of ECGs.

Methods: A total of 63 participants with different levels of clinical expertise took part in an eye-tracking study that consisted of interpreting 10 ECGs with different cardiac abnormalities. A counterbalanced within-subjects design was used with one independent variable consisting of the expertise level of the cardiology practitioners and two dependent variables of eye-tracking metrics (fixations count and fixation revisitations). The eye movements data revealed by specific visual behaviors were analyzed according to the accuracy of interpretation and the frequency with which interpreters visited different parts/leads on a standard 12-lead ECG. In addition, the median and SD in the IQR for the fixations count and the mean and SD for the ECG lead revisitations were calculated.

Results: Accuracy of interpretation ranged between 98% among consultants, 87% among fellows, 70% among technicians, 63% among nurses, and finally 52% among medical students. The results of the eye fixations count, and eye fixation revisitations indicate that the less experienced cardiology practitioners need to interpret several ECG leads more carefully before making any decision. However, more experienced cardiology practitioners rely on their skills to recognize the visual signal patterns of different cardiac abnormalities, providing an accurate ECG interpretation.

Conclusions: The results show that visual expertise for ECG interpretation is linked to the practitioner's role within the health care system and the number of years of practical experience interpreting ECGs. Cardiology practitioners focus on different ECG leads and different waveform abnormalities according to their role in the health care sector and their expertise levels.

(JMIR Hum Factors 2022;9(1):e34058) doi:10.2196/34058

KEYWORDS

eye tracking; electrocardiogram; ECG interpretation; cardiology practitioners; human-computer interaction; cardiology; ECG

Introduction

Visual expertise refers to advanced visual skills demonstrated when executing domain-specific visual tasks [1]. Understanding health care practitioners' visual expertise is crucial in clarifying how to best acquire accurate interpretations of medical images. Visual expertise may be gained through clinical experience, active learning, or teaching. In the context of this study, we shed light on the visual skill of electrocardiogram (ECG) interpretation and, more specifically, on the methodological practices used by cardiology practitioners when conducting visual 12-lead ECG interpretations [2,3]. State-of-the-art visual expertise research has primarily focused on the medical image interpretations of x-rays, mammograms [4,5], and computed tomography and magnetic resonance imaging scans [6,7]. This study is one of few others [3,8-11] that explores how visual expertise contributes to the accuracy of ECG interpretation.

The importance of conducting this study stems from the fact that the ECG is one of the most used medical tests in modern medicine, reaching over 300 million ECGs done annually in the United States alone [12]. In addition, accurate interpretation remains a challenge since there appears to be significant erroneous interpretation rates among nurses, residents, and fellows [3]. Another challenge facing ECG interpreters is the variation in the interpretation procedures and guidelines across different regulating bodies and institutions [13]. Thus, there is a need for additional insights to establish better educational and working practices that suit the different expertise levels of cardiology practitioners to acquire the essential skills for accurately interpreting the ECG. Previous studies focusing on visual expertise in ECG interpretation have mainly restricted their emphasis to the visual aspect of interpretation. Those studies focus primarily on generating eye movement heat maps and statistical data [3,8-11]. However, those same studies lack a discussion on the link between the observational visual behavior of the interpreter and the ECG diagnosis strategy, a critical element that contributes to an accurate ECG interpretation [3].

This study extends the results of our initial work [14-19] under the theme of where does the use of technology fall in the medical landscape. More precisely, one of our studies [14] is aimed at understanding how medical students start to acquire the skill of ECG interpretation. This study focuses more on the cardiology practitioners who interpret ECGs as part of their daily clinical practice. The essence of this study is to pave the way for understanding the link between observational visual behavior and final ECG diagnostics as an element of visual expertise for ECG interpretation. The objective of our study is to quantify and analyze, using eye-tracking technology, differences in cardiology practitioners' visual expertise in ECG interpretation. This quantification is done considering the number of years of practitioners' clinical experience as they advance their medical careers. To reach this objective, we identify eye-tracking metrics that serve this purpose and provide insights into interpretation methodological strategies underpinning accurate ECG interpretation. We then conduct an eye-tracking study with five different categories of cardiology practitioners with different expertise levels. The quantitative results provide insights into interpretation methodological strategies underpinning accurate ECG interpretation, which varies according to the number of years of practical experience in ECG interpretation. Finally, ECG interpretation trends among the pool of participants are unveiled by creating matches between eye fixation heat maps and other eye-tracking metrics.

Methods

Hypotheses

Related works focusing on the relationship between visual expertise, ECG interpretation [3,8-11], and other clinical fields [4-7] requiring medical images interpretation were taken into consideration before creating the following hypotheses. The eye-tracking study by Davies et al [10] especially inspired our second and third hypothesis as the authors noted that experienced interpreters adopt a duel processing model of ECG interpretation. Additionally, the study by Wu et al [3] also emphasized this nuance among different categories of medical practitioners. The following are our three hypotheses:

1. There exists a significant quantifiable difference in the accuracy of the interpretation of each expertise level category of participants as they gain more years of experience.
2. There is a significant correlation between the number of years of participant's experience, depicted by their cardiology practitioner roles, and their fixations' behavior around specific areas of the ECG, demonstrated by the fixations count.
3. There is a significant correlation between the number of years of participant's experience, depicted by their cardiology practitioner roles, and their eye movement transition frequency between different parts/leads on the standard 12-lead ECG, demonstrated by fixation revisitations.

Study Design

The conducted study uses eye-tracking technology to quantify and understand differences in human visual behavior during ECG interpretation. With different clinical roles in the health care sector, recruited participants were tasked with interpreting 10 ECGs with different types of cardiac abnormalities. During their interpretation, their eye movements were recorded using an eye tracker and the collected eye movements data was analyzed quantitatively. Participants were also tasked with selecting their final diagnosis for each ECG from among four available choices or writing down a diagnosis other than those proposed. The choices for each ECG are available in Multimedia Appendix 1.

The experiment used a counterbalanced within-subjects design with the following one independent variable: the expertise level of the cardiology practitioner. This can be quantified as a categorical variable based on the number of years of ECG interpretation experience, as described in Table 1. Medical practitioners may also be placed in one of the clinical categories/roles highlighted in Table 1.

Two measured dependent eye-tracking variables were expected to change when the independent variable changed. These two variables are measured according to our definition of grid-based areas of interest (AOIs). A sample grid-based AOI applied to the normal sinus rhythm ECG can be referred to in Multimedia Appendix 2. Our explanation behind our choice for the grid-based AOIs can be found in our previous work [14]. The following are the two dependent variables:

- The average fixations count for each ECG lead for each category of participant
- The average fixation revisitations for each ECG lead for each category of participant

The experiment also had one control variable. The time given for each participant to look at each ECG was limited to 30 seconds. This time limit allowed for all categories of participants to be held to the same standards in terms of the amount of time given for them to analyze each case. This time limit was chosen by consulting the cardiology consultant and professor involved in designing this experiment. The time is also supported by studies investigating the choice of this parameter within different categories of medical practitioners such as medical students and consultants. The time allowed for scanning an ECG was found to have no statistically significant effect on the result of the diagnosis [20]. It was also found that there is a negative correlation between the duration spent looking at an ECG and the accuracy of the final interpretation provided [21].

Table 1. Corresponding variables to each hypothesis.

Hypothesis	Independent variable	Dependent variable
Hypothesis 1	Years of experience	Accuracy of interpretation
Hypothesis 2	Years of experience	Fixations count
Hypothesis 3	Years of experience	Fixation revisitations

Participants

Table 1 summarizes the demographics of the participants included in this study. A total of 63 participants with varying ECG interpretation expertise were recruited from a university campus and a cardiac hospital. Participants were recruited based on their medical category represented by their job title/role in clinical practice. The mean age was 28 (SD 4) years. In addition, participants were asked to provide an approximation of their years of work experience in ECG interpretation. The medical categories are defined as follows:

- Junior medical students: those in a preclinical curriculum
- Senior medical students: those in a clinical curriculum
- Nurses: nurses either serving in the catheterization laboratory or the cardiac care unit
- Technicians: cardiovascular technologists working in a cardiac catheterization laboratory
- Fellows: physicians undergoing postgraduate training in cardiology
- Cardiology consultants: board-certified independent cardiology practitioners

Stimuli Design

The ECG stimuli were acquired from the collection belonging to the cardiology consultant involved in designing the experiment. Since the study is motivated by quantifying visual behavior across different expertise levels of different health care practitioners, we selected ECGs commonly encountered by all those categories in their day-to-day medical practice [22]. The ECGs sampled are defined in Multimedia Appendix 3. We limited our selection to 10 representative ECG cases.

Apparatus

A Tobii Pro X2-60 eye tracker and iMotions version 8.1 software [23] were used to track eye movements with a frequency of 60 Hz (±1 Hz). In addition, keyboard presses and mouse input were recorded to register the participants' responses showing their final diagnosis for each ECG. The study was conducted on a 25-inch diagonal laptop monitor with a resolution of 1366 by 768 pixels.

Ethics

This study received institutional review board approval from the ethical board of both the Qatar Biomedical Research Institute at Hamad bin Khalifa University [24] under the research protocol number QBRI-IRB-2020-01-009 and the Hamad Medical Corporation under the research protocol number MRC-02-20-714. Approvals were granted before the start of the experiment. Institutional review board approval guarantees that all study methods were conducted following the guidelines and recommendations of international regulatory agencies [24].

Analysis

Hypotheses Testing Methods

The three hypotheses regarding participants' visual behavior toward ECG interpretation were tested as follows.

Analysis Method for Testing Hypothesis 1

To test the first hypothesis, interpretations were assessed for participants' accuracy by determining if they chose the correct ECG diagnosis from among the four offered choices. Analyzing the participants' accuracy of interpretation scores using the Cramér V statistical test contributed toward constructing a clear understanding of how much the interpreters understood the ECG

signals and its waveform abnormalities presented to them throughout the 10 ECG cases.

Analysis Method for Testing Hypothesis 2

To test the second hypothesis, interpretations were assessed for the frequency with which the participants fixated on ECG images. This assessment was done by comparing eye movement behavior for the five categories. Eye movement was quantified using a median fixations count for each participant. A prior study showed that the average duration for one fixation ranged from 150 to 300 milliseconds [25]. Although the average fixation duration span has a fixed range, the fixation count provides a more accurate depiction of the interpreter's attention. The fixations count number represents the interpreter's engagement with different ECG leads suggesting that the greater the median fixation duration, the greater the level of engagement [3].

Analysis Method for Testing Hypothesis 3

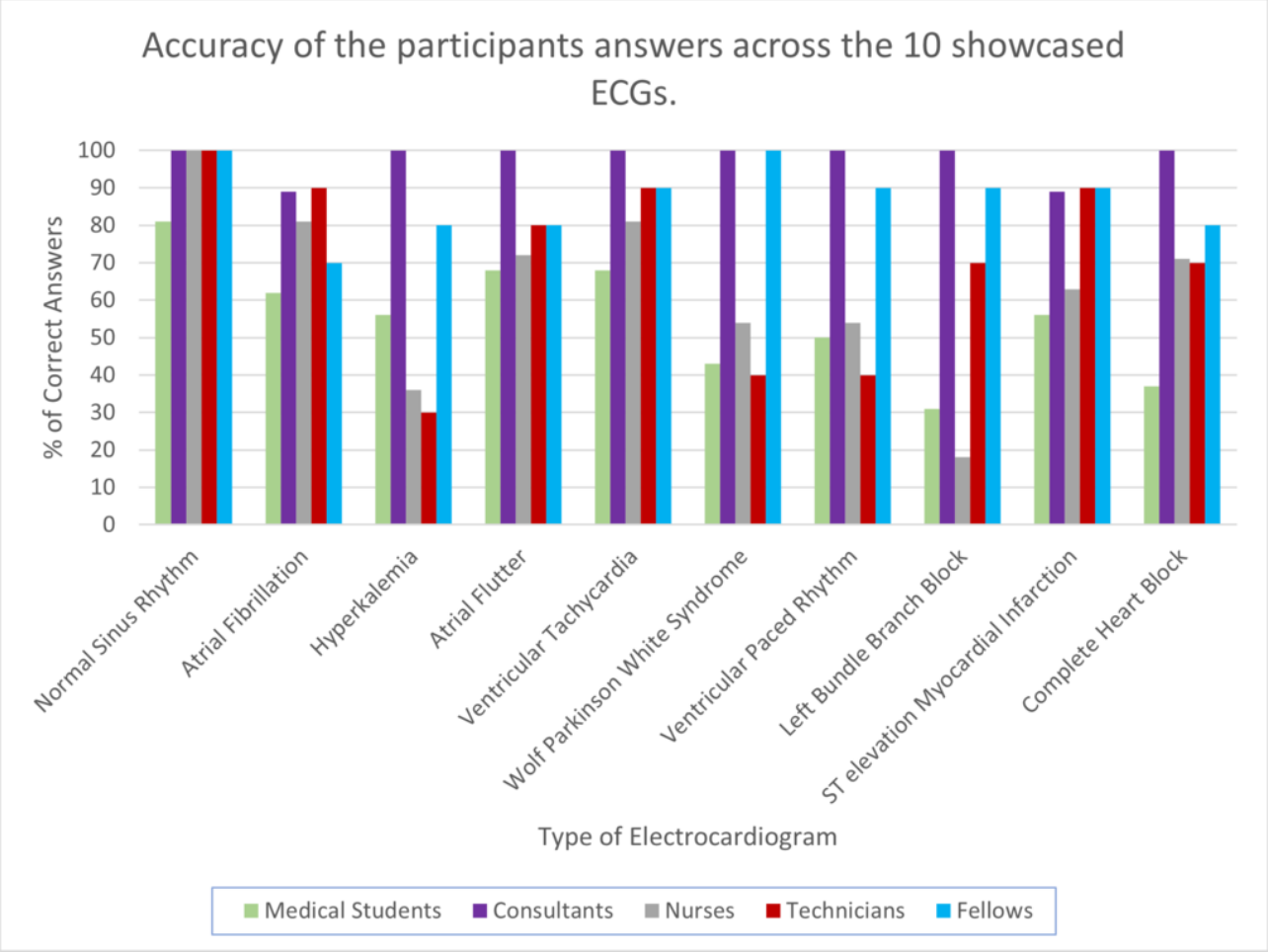
To test the third hypothesis, interpretations were assessed for the frequency with which the interpreters revisit different areas, or leads, in the ECG. This was done by comparing each participant's average ECG lead revisitation among the five categories. A *lead revisitation* is defined as the interpreter fixating again on a particular lead after visiting it previously.

Results

Results for Testing Hypothesis 1

Figure 1 summarizes the accuracy of the participants' answers across the 10 showcased ECGs. Consultants are the most accurate interpreters, with an accuracy percentage of 97.8%. Fellows are the second most accurate, with an overall accuracy of 87%, followed by technicians with an accuracy of 70%. Nurses were the least accurate of those with working experience, with an accuracy of 63%. Finally, medical students were overall the least accurate category with an accuracy of 52.2%. A chi-square test was conducted since the interpretation response data is dichotomous. The obtained *P* value was .02, which shows that there is a statistically significant difference in interpretation accuracy between the five categories. To calculate the effect size of the chi-square independence test, we used the Cramér V, providing a value of 0.36 that indicates a weak association between the categories. Although some ECGs were easy to interpret correctly (eg, the normal sinus rhythm), other ECGs, including the ventricular paced rhythm and the left bundle branch block, were harder to interpret correctly.

Figure 1. Accuracy of the participants' answers across the 10 showcased ECGs. ECG: electrocardiogram.



Results for Testing Hypothesis 2

Table 2 summarizes the median and SD from the IQR for the total fixations count per participant for all leads in the ECG. The fixations data follows a nonnormal, left-skewed distribution. Total fixation counts were calculated for the 10 ECGs for 300 seconds. Table 2 also includes the median and SD from the IQR for the fixation count per lead across all 10 ECGs. Consultants are the category with the lowest number of fixations, while medical students have the highest number of fixations. Applying the Kruskal-Wallis test to the total fixations count data resulted in a *P* value of .03, which showed there is a statistically significant difference in participants' fixation count.

Table 2. Demographics for the participants included in the eye-tracking study.

Feature and demographics	Participants, n
Medical category	
Medical students (junior year)	9
Medical students (senior year)	10
Fellows	11
Technicians	10
Nurses	14
Consultants	9
Age (years)	
20-23	10
23-25	9
26-30	21
30-35	11
35-45	12
Gender	
Male	51
Female	12
Years of experience	
0 years	10
1 year	9
2-5 years	15
5-10 years	22
≥15 years	7

Results for Testing Hypothesis 3

Table 3 summarizes the median total fixations count per participant and median fixation count per lead per participant, while Table 4 summarizes the mean and SD of the ECG lead revisitations for each category of participants. The ECG lead revisitations data follows a normal distribution. On average, technicians are the category of participants with the highest number of revisitations for each lead with an average of 3.61 revisitations, while consultants are the category that revisits a lead the least with an average of 2.01 revisitations. However, based on the SD results, variation among participants in the same category was the highest among nurses and the lowest among technicians. A one-way analysis of variance test was applied to the data, showing an *F* value of 30.56, which is larger than the critical *F* value (2.36). We measured the effect size using the Eta squared formula and the result is η^2=0.36.

Table 3. Median total fixations count per participant and median fixation count per lead per participant.

Category	Fixations count per participant (for all ECGs[a])		Fixation count per lead for each ECG	
	Median	SD from the IQR	Median	SD from the IQR
Medical students	2829	1411	9.93	5.01
Technicians	2535	301	10.83	1.27
Nurses	2444	1031	9.49	3.96
Fellows	2135	579	9.12	2.52
Consultants	1385	794	6.57	3.95

[a]ECG: electrocardiogram.

Table 4. Average electrocardiogram (ECG) lead revisitation per participant for every category.

Category	ECG lead revisitation per participant	
	Mean (μ)	SD (σ)
Technician	3.61	0.06
Nurse	3.25	1.60
Medical students	2.90	0.85
Fellow	2.55	0.67
Consultant	2.01	0.98

Discussion

Insights From the Eye-Tracking Results

The results indicate that the interpreter's expertise, revealed by the number of years of work experience in ECG interpretation, is the primary influence for both the accuracy of ECG interpretation and the acquired visual expertise strategies. Through the analysis of the three hypotheses, three main findings were confirmed.

First, the accuracy of ECG interpretation correlates with the expertise level of the participant. The results confirm the first hypothesis by indicating that consultants are the category with the most accurate interpretations, while medical students are the category with the least accurate interpretations. In between these two extremes are nurses, technicians, and fellows.

Second, as expertise for participants increases, participants' fixations count on ECG signal waveform abnormalities decreases. This finding translates into participants fixating on the overall ECG for less time while not compromising the accuracy of the interpretation. This finding confirms the second hypothesis.

Third, the results for testing hypothesis 3 confirm that medical practitioners observe and focus on certain ECG leads and waveform abnormalities according to their role in the health care sector and their expertise level. Figures 2 and 3 show sample aggregate heat maps for the differences in fixation distribution across the left bundle branch block and the complete heart block ECGs between the studied categories. The results from the third assessment aiming to confirm the third hypothesis indicate that both consultants and fellows target their fixations on specific leads to identify the correct ECG diagnosis. This finding was also confirmed by looking at the heat maps for different categories across different ECGs other than the ones in Figures 2 and 3. However, nurses and technicians thoroughly interpreted the ECGs systematically by scanning through all 12 leads and primarily looking at abnormalities in the ST segment and wide QRS complex. This finding explains the high number of fixations per lead measured by the fixations count. This behavior may be because nurses are usually not extensively trained to interpret the ECG the way cardiologists do for a diagnosis but to ensure that signs of imminent heart attacks are not missed. The heat maps of medical students indicate that they randomly fixate on the waveform abnormalities that they first perceive and then continue to transition from one lead to another until the 30-second time limit is over. This finding justifies medical students having the highest number of fixations per ECG case. Heat maps for all the categories of interpreters and all the ECGs can be found in Multimedia Appendix 4.

Figure 2. Sample aggregate heat maps showing differences in fixation distribution across the left bundle branch block ECG between the studied categories. ECG: electrocardiogram.

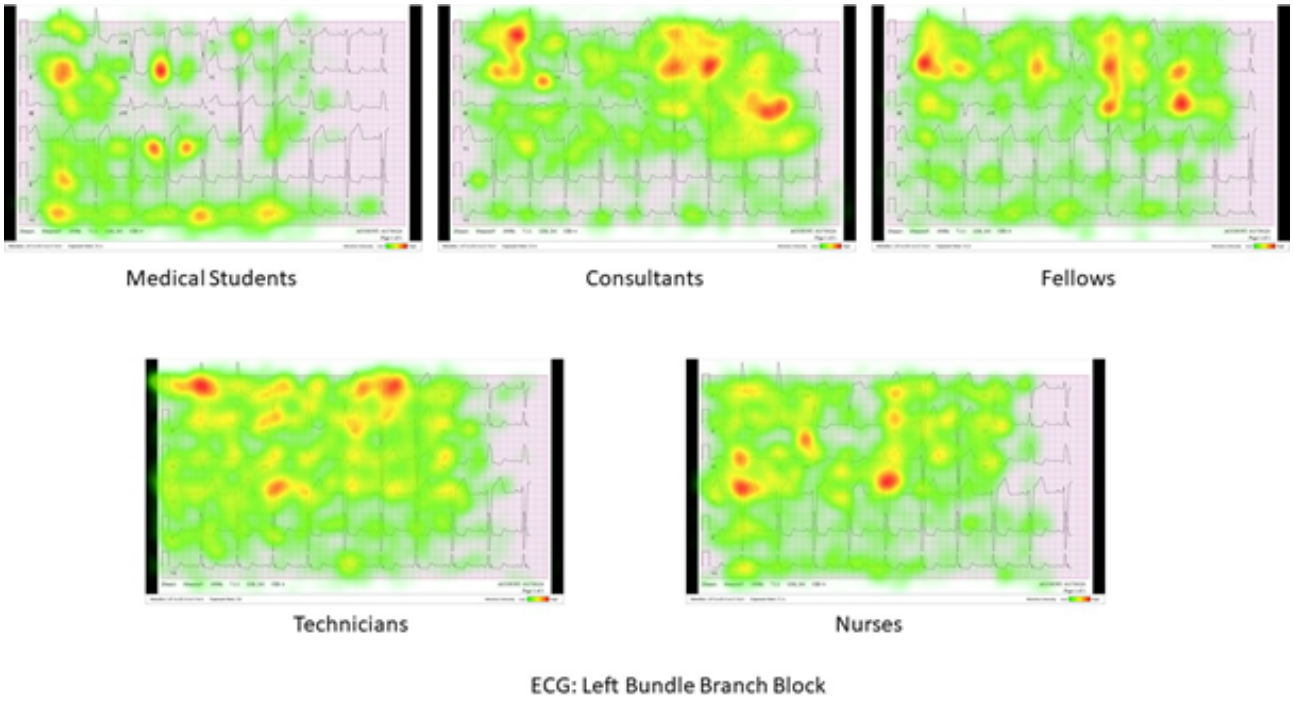
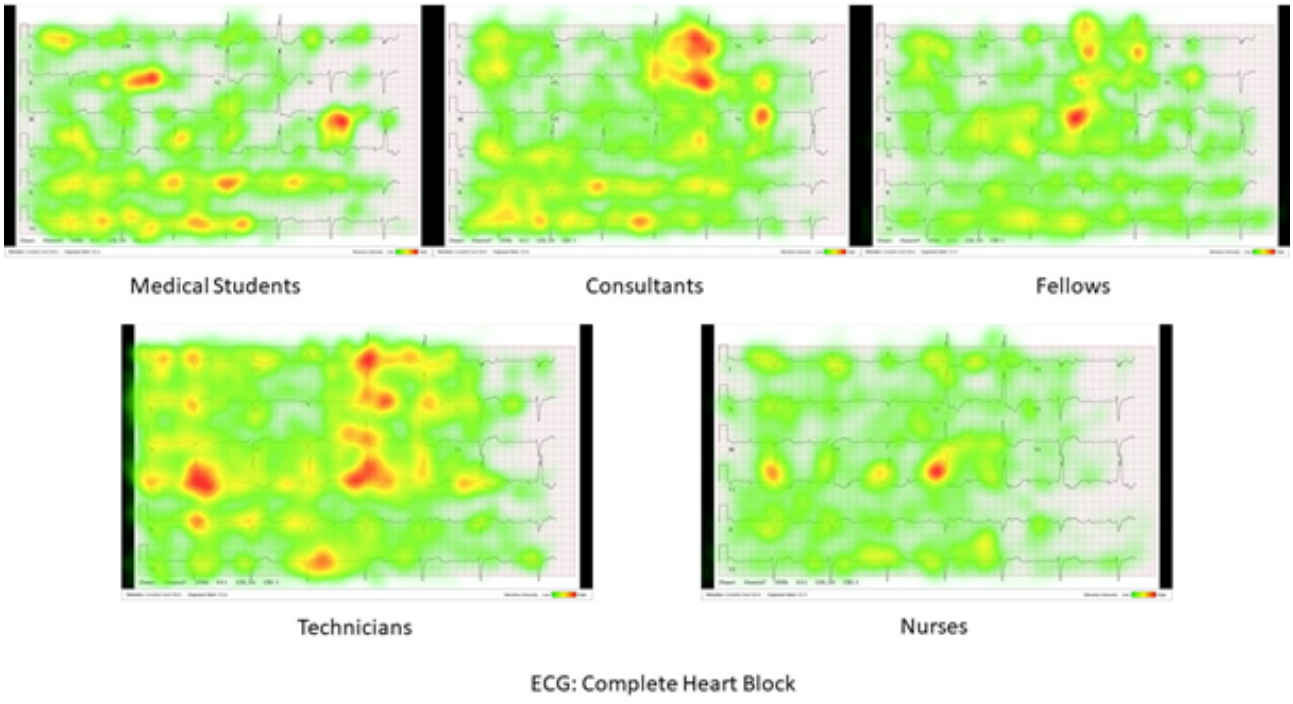


Figure 3. Sample aggregate heat maps showing differences in fixation distribution across the complete heart block ECG between the studied categories. ECG: electrocardiogram.



Conclusion

Concluding Remarks

This paper presents a quantitative analysis of the ECG interpretation visual behavior of different health care practitioners. These health care practitioners belonged to five different categories: medical students, nurses, technicians, fellows, and consultants. Eye-tracking data for these categories were collected while they each interpreted a total of 10 ECGs. Specific eye-tracking metrics such as fixations count and fixation revisitations were quantitatively analyzed for each lead in the standard 12-lead ECG and across all ECGs. This analysis was done to meet the objective of quantifying, using eye tracking, medical practitioners' visual expertise strategies in ECG interpretation as they advance in their medical careers. The main findings relate to how accurate each medical category

is in ECG interpretation when considering their eye movements and visual behavior. The accuracy of the final ECG diagnosis was also associated with the expertise level of participants. Moreover, the increased level of participant expertise made experienced practitioners require less time to fixate on ECG abnormalities and decreased fixation counts, leading to correct diagnoses. Lastly, medical practitioners focus on certain ECG leads and specific waveform abnormalities according to their role in the health care sector and their expertise level.

Study Limitations and Future Works

Since eye-tracking data is idiosyncratic to every interpreter, a sample size of approximately 60 participants from different categories may not be representative enough. Sample size determination depends on what the designers of the study aim to represent. Recruited sample size may therefore vary according to the targeted population, CIs, and interpreters' confidence level in their responses. Based on these uncontrollable factors, recruiting more participants and increasing the number of medical practitioner categories are necessary. We addressed this by recruiting a diverse and reasonable number of health care practitioners, but including larger numbers of participants in future work would contribute to a better understanding of visual expertise in ECG interpretation and understanding how different health care practitioners with different roles and expertise levels interpret ECGs. The richness of the study's collected eye movement data has the potential to be further analyzed using machine learning algorithms to deeply reveal differences in visual behavior among the different categories of medical practitioners. We also plan on experimenting with more subtle examples of ECG diagnoses such as nonspecific/incomplete abnormalities and see how the experts deal with conflicting or vague data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Multiple choice questions for the electrocardiogram eye-tracking experiment.
[PDF File (Adobe PDF File), 74 KB - humanfactors_v9i1e34058_app1.pdf]

Multimedia Appendix 2
Sample grid-based areas of interest applied to the normal sinus rhythm electrocardiogram.
[PNG File , 1318 KB - humanfactors_v9i1e34058_app2.png]

Multimedia Appendix 3
Definition of electrocardiogram samples used in the eye-tracking experiment.
[PDF File (Adobe PDF File), 93 KB - humanfactors_v9i1e34058_app3.pdf]

Multimedia Appendix 4
Heat maps for all the categories of interpreters and all the electrocardiograms.
[RAR File , 37173 KB - humanfactors_v9i1e34058_app4.rar]

References

1. Gegenfurtner A, van Merriënboer JJG. Methodologies for studying visual expertise. Frontline Learning Res 2017 Jul 14;5(3):1-13. [doi: 10.14786/flr.v5i3.316]
2. Heiberg Engel PJ. Tacit knowledge and visual expertise in medical diagnostic reasoning: implications for medical education. Med Teach 2008;30(7):e184-e188. [doi: 10.1080/01421590802144260] [Medline: 18777417]
3. Wu W, Hall AK, Braund H, Bell CR, Szulewski A. The development of visual expertise in ECG interpretation: an eye-tracking augmented re situ interview approach. Teach Learn Med 2021;33(3):258-269. [doi: 10.1080/10401334.2020.1844009] [Medline: 33302734]
4. Kundel HL, Nodine CF, Conant EF, Weinstein SP. Holistic component of image perception in mammogram interpretation: gaze-tracking study. Radiology 2007 Feb;242(2):396-402. [doi: 10.1148/radiol.2422051997] [Medline: 17255410]
5. Mello-Thoms C, Dunn S, Nodine C, Kundel H. The perception of breast cancers--a spatial frequency analysis of what differentiates missed from reported cancers. IEEE Trans Med Imaging 2003 Oct;22(10):1297-1306. [doi: 10.1109/TMI.2003.817784] [Medline: 14552583]
6. Manning D, Ethell S, Donovan T, Crawford T. How do radiologists do it? The influence of experience and training on searching for chest nodules. Radiography 2006 May;12(2):134-142. [doi: 10.1016/j.radi.2005.02.003]
7. Kundel HL, Nodine CF. Interpreting chest radiographs without visual search. Radiology 1975 Sep;116(3):527-532. [doi: 10.1148/116.3.527] [Medline: 125436]

8. Wood G, Batt J, Appelboam A, Harris A, Wilson MR. Exploring the impact of expertise, clinical history, and visual search on electrocardiogram interpretation. Med Decis Making 2014 Jan;34(1):75-83. [doi: 10.1177/0272989X13492016] [Medline: 23811761]
9. Bond R, Kligfield P, Zhu T, Finlay D, Drew B, Guldenring D, et al. Novel approach to documenting expert ECG interpretation using eye tracking technology: a historical and biographical representation of the late Dr Rory Childers in action. J Electrocardiol 2015;48(1):43-44. [doi: 10.1016/j.jelectrocard.2014.10.017] [Medline: 25465863]
10. Davies A, Mueller J, Horseman L, Splendiani B, Hill E, Vigo M, et al. How do healthcare practitioners read electrocardiograms? A dual-process model of electrocardiogram interpretation. Br J Cardiac Nurs 2019 Oct 01;14(10):1-19. [doi: 10.12968/bjca.2019.0073]
11. Breen CJ, Bond R, Finlay D. An evaluation of eye tracking technology in the assessment of 12 lead electrocardiography interpretation. J Electrocardiol 2014;47(6):922-929. [doi: 10.1016/j.jelectrocard.2014.08.008] [Medline: 25200901]
12. Cadet JV. Report: cost savings will drive ECG global market to nearly $160M by 2015. Cardiovascular Business. 2009. URL: https://www.cardiovascularbusiness.com/topics/cardiovascular-imaging/report-cost-savings-will-drive-ecg-global-market-nearly-160m-2015 [accessed 2020-08-10]
13. Kligfield P, Gettes LS, Bailey JJ, Childers R, Deal BJ, Hancock EW, American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology, American College of Cardiology Foundation, Heart Rhythm Society, et al. Recommendations for the standardization and interpretation of the electrocardiogram: part I: The electrocardiogram and its technology: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society: endorsed by the International Society for Computerized Electrocardiology. Circulation 2007 Mar 13;115(10):1306-1324. [doi: 10.1161/CIRCULATIONAHA.106.180200] [Medline: 17322457]
14. Tahri Sqalli M, Al-Thani D, Elshazly MB, Al-Hijji M. Interpretation of a 12-lead electrocardiogram by medical students: quantitative eye-tracking approach. JMIR Med Educ 2021 Oct 14;7(4):e26675 [FREE Full text] [doi: 10.2196/26675] [Medline: 34647899]
15. Tahri Sqalli M, Al-Thani D. AI-supported health coaching model for patients with chronic diseases. 2019 Presented at: 16th International Symposium on Wireless Communication Systems; August 27-30, 2019; Oulu, Finland. [doi: 10.1109/iswcs.2019.8877113]
16. Tahri Sqalli M, Al-Thani D, Qaraqe M, Fernandez-Luque L. Perspectives on human-AI interaction applied to health and wellness management: between milestones and hurdles. In: Househ M, Borycki E, Kushniruk A, editors. Multiple Perspectives on Artificial Intelligence in Healthcare: Opportunities and Challenges. Cham: Springer; 2021:41-51.
17. Tahri Sqalli M, Al-Thani D. Towards an integrated health and wellness program using human-IoT interaction. 2020 Presented at: 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies; February 2-5, 2020; Doha, Qatar. [doi: 10.1109/iciot48696.2020.9089538]
18. Tahri Sqalli M, Al-Thani D. On how chronic conditions affect the patient-AI interaction: a literature review. Healthcare (Basel) 2020 Sep 01;8(3):313 [FREE Full text] [doi: 10.3390/healthcare8030313] [Medline: 32883036]
19. Tahri Sqalli M, Al-Thani D. Evolution of wearable devices in health coaching: challenges and opportunities. Front Digit Health 2020;2:545646 [FREE Full text] [doi: 10.3389/fdgth.2020.545646] [Medline: 34713031]
20. Davies AR. Examining expertise through eye movements: a study of clinicians interpreting electrocardiograms. The University of Manchester Library. 2018. URL: https://www.escholar.manchester.ac.uk/api/datastream?publicationPid=uk-ac-man-scw:314626&datastreamId=FULL-TEXT.PDF [accessed 2022-02-03]
21. Bond R, Zhu T, Finlay D, Drew B, Kligfield P, Guldenring D, et al. Assessing computerized eye tracking technology for gaining insight into expert interpretation of the 12-lead electrocardiogram: an objective quantitative approach. J Electrocardiol 2014;47(6):895-906. [doi: 10.1016/j.jelectrocard.2014.07.011] [Medline: 25110276]
22. Davies A, Harper S, Vigo M, Jay C. Investigating the effect of clinical history before electrocardiogram interpretation on the visual behavior and interpretation accuracy of clinicians. Sci Rep 2019 Aug 05;9(1):11300. [doi: 10.1038/s41598-019-47830-0] [Medline: 31383896]
23. Farnsworth B. 10 most used eye tracking metrics and terms. iMotions. 2020. URL: https://imotions.com/blog/10-terms-metrics-eye-tracking/ [accessed 2022-02-03]
24. Research Compliance Assurance. Hamad Bin Khalifa University. URL: https://www.hbku.edu.qa/en/qbri/research-areas/research-compliance [accessed 2022-02-03]
25. Tullis T, Albert B. Behavioral and physiological metrics. In: Measuring the User Experience Collecting, Analyzing, and Presenting Usability Metrics. Amsterdam: Elsevier Inc; 2013:163-186.

Abbreviations

AOI: area of interest
ECG: electrocardiogram

Edited by A Kushniruk; submitted 05.10.21; peer-reviewed by A Davies, B Walsh, N Silva; comments to author 28.11.21; revised version received 08.12.21; accepted 09.12.21; published 09.02.22.

Please cite as:
Tahri Sqalli M, Al-Thani D, Elshazly MB, Al-Hijji M, Alahmadi A, Sqalli Houssaini Y
Understanding Cardiology Practitioners' Interpretations of Electrocardiograms: An Eye-Tracking Study
JMIR Hum Factors 2022;9(1):e34058
URL: https://humanfactors.jmir.org/2022/1/e34058
doi:10.2196/34058
PMID:35138258

Original Paper

User-Centered Design of A Novel Risk Prediction Behavior Change Tool Augmented With an Artificial Intelligence Engine (MyDiabetesIQ): A Sociotechnical Systems Approach

Cathy Shields[1], BSc, PhD; Scott G Cunningham[1], BSc, PhD; Deborah J Wake[2], MBChB (hons), BSc, PhD, DipMedEd; Evridiki Fioratou[3], BSc, PhD; Doogie Brodie[4]; Sam Philip[5], MBBS, MD, MRCP; Nicholas T Conway[6], MBChB, MRCPCH, MPH, MD

[1]Division of Population Health and Genomics, School of Medicine, University of Dundee, Dundee, United Kingdom

[2]Usher Institute, University of Edinburgh, Edinburgh, United Kingdom

[3]Ninewells Hospital and Medical School, University of Dundee, Dundee, United Kingdom

[4]MyWay Digital Health, Dundee, United Kingdom

[5]Grampian Diabetes Research Unit, Aberdeen Royal Infirmary, Aberdeen, United Kingdom

[6]NHS Tayside, Dundee, United Kingdom

Corresponding Author:
Cathy Shields, BSc, PhD
Division of Population Health and Genomics
School of Medicine
University of Dundee
Mackenzie Building
Kirsty Semple Way
Dundee, DD2 4BF
United Kingdom
Phone: 44 01382 381382
Email: c.v.shields@dundee.ac.uk

Abstract

Background: Diabetes and its complications account for 10% of annual health care spending in the United Kingdom. Digital health care interventions (DHIs) can provide scalable care, fostering diabetes self-management and reducing the risk of complications. Tailorability (providing personalized interventions) and usability are key to DHI engagement/effectiveness. User-centered design of DHIs (aligning features to end users' needs) can generate more usable interventions, avoiding unintended consequences and improving user engagement.

Objective: MyDiabetesIQ (MDIQ) is an artificial intelligence engine intended to predict users' diabetes complications risk. It will underpin a user interface in which users will alter lifestyle parameters to see the impact on their future risks. MDIQ will link to an existing DHI, My Diabetes My Way (MDMW). We describe the user-centered design of the user interface of MDIQ as informed by human factors engineering.

Methods: Current users of MDMW were invited to take part in focus groups to gather their insights about users being shown their likelihood of developing diabetes-related complications and any risks they perceived from using MDIQ. Findings from focus groups informed the development of a prototype MDIQ interface, which was then user-tested through the "think aloud" method, in which users speak aloud about their thoughts/impressions while performing prescribed tasks. Focus group and think aloud transcripts were analyzed thematically, using a combination of inductive and deductive analysis. For think aloud data, a sociotechnical model was used as a framework for thematic analysis.

Results: Focus group participants (n=8) felt that some users could become anxious when shown their future complications risks. They highlighted the importance of easy navigation, jargon avoidance, and the use of positive/encouraging language. User testing of the prototype site through think aloud sessions (n=7) highlighted several usability issues. Issues included confusing visual cues and confusion over whether user-updated information fed back to health care teams. Some issues could be compounded for users with limited digital skills. Results from the focus groups and think aloud workshops were used in the development of a live MDIQ platform.

Conclusions: Acting on the input of end users at each iterative stage of a digital tool's development can help to prioritize users throughout the design process, ensuring the alignment of DHI features with user needs. The use of the sociotechnical framework encouraged the consideration of interactions between different sociotechnical dimensions in finding solutions to issues, for example, avoiding the exclusion of users with limited digital skills. Based on user feedback, the tool could scaffold good goal setting, allowing users to balance their palatable future complications risk against acceptable lifestyle changes. Optimal control of diabetes relies heavily on self-management. Tools such as MDMW/ MDIQ can offer personalized support for self-management alongside access to users' electronic health records, potentially helping to delay or reduce long-term complications, thereby providing significant reductions in health care costs.

(JMIR Hum Factors 2022;9(1):e29973) doi:10.2196/29973

KEYWORDS

diabetes mellitus; digital health intervention; eHealth; artificial intelligence; user-centred design; human factors; think aloud

Introduction

Diabetes is a chronic disease affecting an estimated 463 million adults worldwide [1], with global spending related to diabetes and its complications exceeding US $800 billion annually [2]. Good control of diabetes decreases the risk of associated chronic complications [3]. Digital health interventions (DHIs), delivered through interactive websites and mobile apps, have the potential to harness the omnipresence and growing computational power of electronic devices for the self-management of chronic diseases [4-6]. Usability and "tailorability" are known to improve user acceptability and engagement with DHIs [7,8] and, therefore, potentially, their effectiveness. A relatively small number of studies have so far been published on DHI usability [9], partly due to many such interventions being developed by commercial companies [10]. Fewer still have sought to include "end users" in usability testing.

The UK government's "Five Year Forward View" encourages and enables individuals to take greater responsibility for their health through the use of eHealth or mobile health services [11], a sentiment that is mirrored in most high-income nations. This drive has become more pressing due to the impact of COVID-19, which has altered care models for people with diabetes (PWD) [12], necessarily shifting the focus towards remote care [13]. Among all chronic conditions, diabetes is possibly the most amenable to the use of DHIs in scalable follow-up care [14,15]. Many of the challenges faced by people in managing their diabetes occur "in the moments of everyday life" [15] when a DHI has the potential to deliver targeted and timely assistance.

Considering how crucial usability is for a DHI, it is concerning that a recent evaluation of four top-rated diabetes apps found that they all suffered from usability problems, many of which were "very serious" [16]. User-centered design of health interventions explicitly involves end users in their design, development, and evaluation [14,17]. This approach has the potential to produce more acceptable and usable DHIs [18] by ensuring from the outset that an intervention is targeted to end users' needs [19,20]. The use of iterative design cycles, which include end users at each stage of product development, is also recognized as being important in developing usable DHIs [10]. Gaining a good understanding of the context in which specified users will interact with the product is also important [21].

Sittig and Singh's sociotechnical model [22] was developed to allow the social context of digital health care tools to be linked to the technical component, and it recognizes that the two components influence one another [23]. It has previously been adapted by others to examine a range of different health care technologies, including patient-facing portals and health apps [24-26]. It allows for different sociotechnical dimensions to be dismantled for the purposes of examining them but also encourages consideration of the relationships between dimensions [23]. The sociotechnical approach encompasses a human factors engineering approach, which attempts to optimize users' performance of tasks whilst making allowances for human capabilities and limitations in complex environments [27].

In recent times, it has become possible for artificial intelligence (AI), including machine learning (ML) algorithms, to underpin DHIs. ML algorithms analyze large data sets to detect patterns in data [28]. They can generate predictive models, the outputs of which can support decision-making by users [29] and include predictions of future risks. Effective communication of risk within a DHI must take into account end users' health literacy and numeracy [30]. In addition, DHI interfaces should adhere to evidence-based recommendations for the presentation of complex risk information to patients (eg, use of plain language and use of absolute rather than relative risk) [26,30].

My Diabetes My Way (MDMW) [31,32] is a DHI (interactive website and app) for PWD and their carers. It contains resources, including tailored information based on users' health/lifestyle data, interactive educational resources, and videos, as well as access to users' clinical data. MDMW has been used in Scotland since 2008 and has more recently (since 2018) been deployed in several National Health Service (NHS) Trusts in England (ie, Somerset, Northeast London, Lancashire, South Cumbria, Cheshire, and Merseyside), where it is known as MyWay Diabetes. MDMW takes a subset of data from primary and (where possible) secondary care. These include key diabetes indicators (eg, glycated hemoglobin [HbA_{1c}], blood pressure, and BMI), as well as eye and foot screening results, medication, and clinical correspondence. It provides users access to these records, as well as tailored advice and targeted resources based on each user's status. History graphs permit individuals to interrogate their data and progress over time. One area of the site (the "managing your condition" page) alerts users to missed screening visits (based on the Diabetes UK "15 Healthcare Essentials" [11]). Patients can manually enter home-recorded

data (eg, weight, blood pressure, and blood glucose) and set their own health and lifestyle goals.

An AI-augmented version of MDMW is being developed through linkage to the MyDiabetesIQ (MDIQ) analytics and reporting engine. MDIQ is linked to a knowledge base of approximately 70 validated machine learning models. It uses information from health care records (ie, linkages with hospital/general practitioner information technology systems) and home recordings (eg, Fitbit activity and home blood glucose data), driving predictive analytics. MDIQ models were developed in two ways: (a) using individual literature-published models and (b) meta-models derived from literature-published models, but delivering improved performance. Models were revalidated and tested in Scottish diabetes data sets, including NHS Greater Glasgow and Clyde (n=105,000) and Northwest London data sets (n=145,000). Proposed novel features underpinned by MDIQ include: (1) presenting users with their predicted risk of diabetes-related complications (based on their clinical and lifestyle data), (2) allowing users to visualize how lifestyle changes could impact their risk, (3) and providing ongoing, tailored feedback on progress toward users' own goals.

This study aims to provide an overview of the iterative design process of the enhanced interface of this DHI and how human factors have informed system development using the sociotechnical model as a theoretical framework [22]. MDIQ AI models underpinning these additional features will be described in greater detail in a separate publication.

Methods

Focus Groups

Potential focus group participants were identified via local patient and public involvement (PPI) groups whose members had previously expressed an interest in taking part in diabetes-related research and who had consented to be contacted. Initial email contact was made by local research nurses. Interested patients were directed to contact the researcher for further information and to complete the consent process. Two of the authors (CS plus one other) moderated the focus group sessions, which were held in meeting rooms within a local private venue or hospital. Sessions were guided by an interview schedule (Multimedia Appendix 1) and screenshots of early designs (for a new-look homepage, goal-setting area, and novel risk prediction tool).

Focus groups were audio-recorded and later transcribed verbatim. NVivo 12 software (QSR International Ltd) was used to organize and code the textual data, which were analyzed using inductive thematic analysis [33]. In an effort to reduce bias, qualitative data were coded by 2 authors, a researcher (CS) and a clinician (NC). Any conflicts in coding decisions were reviewed by both and resolved by discussion and consensus [34].

Where participants are quoted in the text, they are referred to by the number of the focus group they attended and their participant number (eg, FG1P1 = focus group 1, participant 1), along with their sex, age, and diabetes type.

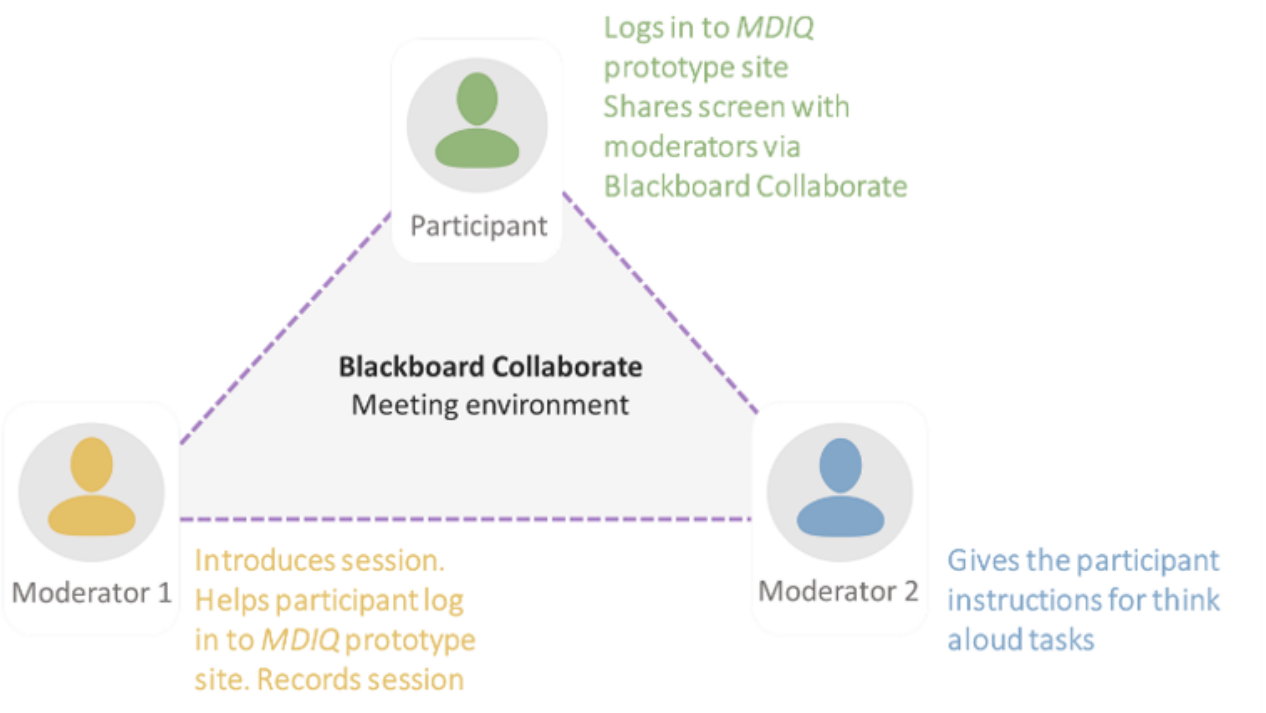
NHS Health Research Authority Research Ethics Committee approvals were obtained (IRAS number: 258231).

Think Aloud Workshops

Focus group participants were invited back to attend workshops to user test a prototype site, the development of which was guided by focus group findings. Not all participants returned; therefore, several additional participants were recruited from the original PPI list. Workshops followed the "think aloud" method [35,36]. In this method, the participant is asked to explain their thinking, opinions, actions, and reactions as they perform several prescribed tasks [9]. Tasks given to users (Multimedia Appendix 2) represented common actions for which the system would eventually be used. Several tasks were similar to actions that participants would be familiar with when using the existing MDMW site. Other tasks were novel, relating to unique MDIQ-augmented features. All tasks were validated by 2 of the authors (CS and NC) and collaborator Louise McIver.

Think aloud sessions took place in August 2020 and were conducted remotely due to COVID-19 restrictions at the time of the sessions. A schematic of the set-up for remote user testing is shown in Figure 1. Each participant was sent a Blackboard Collaborate link several days prior to their session. Blackboard Collaborate was chosen because no special software download was required by participants, who could then easily screen share. Sessions were recorded by the moderators, with participants' screens and speech captured simultaneously. Two moderators were present in each session: one (CS) introduced the activity, recorded the session, and made notes; the other (Louise McIver) gave instructions for think aloud tasks and encouraged the participant to continue to speak if they became quiet for too long.

Figure 1. Schematic showing how think aloud method was performed remotely, using Blackboard Collaborate Ultra for the meeting environment. MDIQ: MyDiabetesIQ.



Video recordings of sessions were later transcribed verbatim, and each transcript was annotated with observations (eg, position/activity of the user's cursor). A combination of deductive and inductive thematic analysis was employed [33]. Text fragments relating to usability problems were mapped onto all relevant dimensions of the sociotechnical model [22] (ie, human-computer interface, clinical content, workflow and communication, people, and hardware and software), which became the major themes. Within each of the dimensions, an inductive approach was used to identify subthemes that were generated from the data. Qualitative data were coded by 2 authors, a researcher (CS) and a clinician (NC). Any conflicts in coding decisions were reviewed by both and resolved by discussion and consensus [34].

Where participants are quoted in the text, they are referred to by participant number (eg, TA_P1), along with their sex, age, and diabetes type.

The think aloud component of the study was deemed to be "service evaluation;" therefore, ethics approval was not required.

Results

Focus Groups

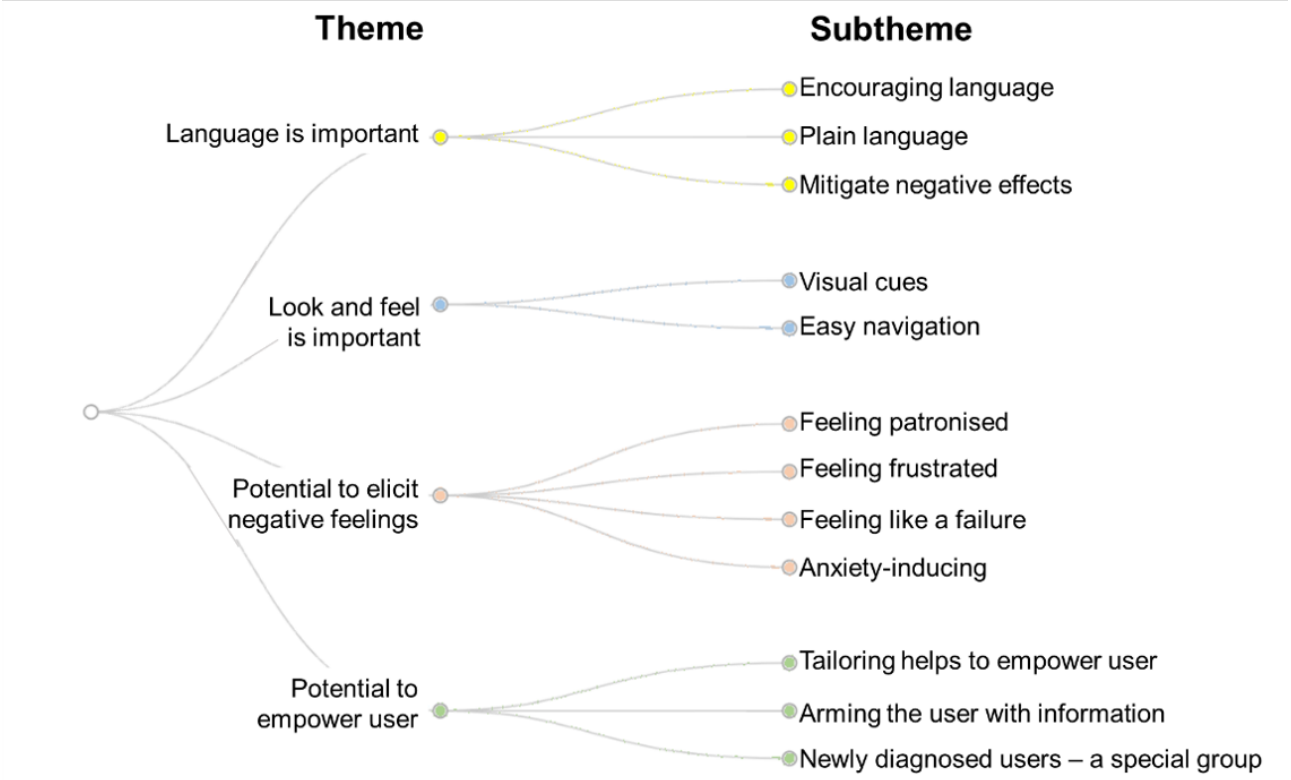
Focus Group Participant Demographics

Two focus groups were conducted in November 2019 and January 2020 in two different regions of Scotland, with a total of 8 participants. There were 2 participants in the first group and 6 in the second group, giving a total of 8 participants. Most participants were male (6/8, 75%), a majority (5/8, 63%) had type 1 diabetes (T1D), and 3 (37%) had type 2 diabetes (T2D). The mean age was 64 years (range 49-83 years), and the mean time since diagnosis was 22 years (range 5-58 years). The average recording length of focus group sessions was 64 min.

Thematic Analysis

Themes and subthemes generated from focus group transcripts through inductive thematic analysis are highlighted in Figure 2. These are discussed below in relation to novel features of the MDIQ-augmented site.

Figure 2. Themes and subthemes identified from focus group discussions.



Potential to Empower Users

The majority of participants (5/8, 63%) felt that some proposed new MDMW features had the potential to empower users. For example, one novel feature described to participants during focus groups was the potential for users to set their own goals via the system and then receive ongoing tailored feedback on progress made toward goals. Two (25%) participants felt that this could help to empower users. An example of this sentiment is given below:

> *You'd set your weight goal and then say 'every week I'm going to step on the scales,' and then you'd put it in. It could be a useful tool as a specific thing. Especially if your doctor has wagged their finger at you.* [FG2P1, male, age 58, T1D]

In the case of a proposed new risk prediction tool, which would display a user's future risk of developing diabetes-related complications and allow them to alter sliders representing lifestyle choices to see how this impacts their risk, 5 (63%) participants felt that this had the potential to empower users by arming them with information, allowing them to take control:

> *You've got the recipient taking decisions and setting their own targets...That might go down well with some folk. Very much so.* [FG1P2, male, age 83, T2D]
>
> *…that's what you want. With diabetes, in particular, you want the patient to have control themselves.* [FG1P2, male, age 83, T2D]

Those participants who were positive about the risk prediction tool saw it as a way of facilitating users setting appropriate goals.

Potential to Elicit Negative Feelings

Despite the acknowledgment that users could be empowered by proposed features, 6 (75%) participants also felt that these could have undesirable effects such as eliciting negative feelings, including users becoming anxious or feeling as though they had "failed." In the case of goal-setting with ongoing feedback, there were concerns about the consequences if a user did not achieve their goals and that this might produce feelings of failure:

> *I'm always wary about giving folk targets—what if they don't make the target? There is a consequence.* [FG1P2, male, age 83, T2D]
>
> *I looked at it last night, and it said, 'your height is this, and your weight is that.' It says, 'you are overweight.' Now I was overweight by 200g! You know? But it said, 'You are overweight. You've failed. You've failed.' And that's down to the language that's in there. (referring to how the current MDMW system reports the user's current weight).* [FG2P1, male, age 58, T1D]

It was pointed out by 3 (38%) participants that goals should be "achievable" and "safe" and that some guidance may be needed on what constitutes a realistic goal. This would help to mitigate against risks:

> *Yes, because they could be setting themselves up for failure just by making it [the goal] unrealistic.* [FG2P3, female, age 62, T1D]
>
> *For those of us who are patients—we should agree the goal with the clinician. My goal is: 'I want to shave down my HbA_{1c} or lose a few pounds or drive down my cholesterol' or whatever it might be. If they*

> *say, 'yes, you could maybe do that in the next 4 months' [sic]. I think it could be dangerous if you could randomly set your own goals.* [FG2P1, male, age 58, T1D]

Although recognized as potentially empowering for users, the risk prediction tool was seen by 4 (50%) participants to have the potential to cause anxiety (or to exacerbate existing health anxieties) due to users' risks being revealed to them:

> *You can have a toggle to say 'I don't want to see that,' but again if someone is worried—is a natural worrier—then they would click 'yes, I want to see that kind of information' and then it's a positive—or a negative—feedback.* [FG1P1, male, age 62, T2D]

> *If folk are prone to worry about things, it could make it worse. But for most of us we're quite pragmatic about it when you've had diabetes for a while.* [FG2P5, female, age 49, T1D]

Newly diagnosed users were regarded by 4 (50%) participants as a special group, who may feel more worried than other users when their future risks are revealed to them. However, it was also felt that they could be the group who could benefit the most from this information:

> *New folk—it might scare them a bit, but…you can't have too much information.* [FG2P5, female, age 49, T1D]

> *For newly diagnosed people, it would be really useful, and I would have found this—yes scary—but still giving me more information.* [FG2P1, male, age 58, T1D]

Language Is Important

In terms of mitigating any anxiety arising from using the risk prediction tool, 5 (63%) participants agreed that the use of positive and encouraging language would be vital:

> *I think you have to say that because it is lifestyle parameters, all of those can be addressed.* [FG2P1, male, age 58, T1D]

> *Yes, so it's not all doom and gloom. This is what it would look like if you carry on, but there is a process for improving on that.* [FG2P2, male, age 66, T1D]

In order to further mitigate against anxiety about complications risk, 3 (38%) participants felt that users should be provided with extra guidance, such as information or a link alongside the risk prediction tool, directing them to mental health advice or suggesting that they speak to their health care provider (HCP) if they are concerned about what they have seen:

> *People choose whether to see that or not, but if they do choose to see it, and the results are bad, then you could maybe put a link in or something to take them to the mental health questionnaire.* [FG1P1, male, age 62, T2D]

> *How difficult would it be [for the site developers] if you're on it and you want to click on a link for support, like your diabetic nurse or your clinic?* [FG2P6, male, age 64, T2D]

"Look and Feel" of the Site

The "look and feel" of the site was also considered important by 4 (50%) of the participants. Discussions on this theme fell into two subthemes: (1) that images and icons would be preferable to large blocks of text (or lists of data) and (2) that navigation should be easy and intuitive:

> *…you end up with a string of menus. You look for your own measurements—there's eight or nine of them, and you work your way down them, and ok by the time you get to number nine you're beginning to forget what was number one! (referring to current MDMW site)* [FG1P2, male, age 83, T2D]

> *If you were going into the front page with your five circles, and you tapped into one of the reds, and you got your graph, that would make sense. (responding to screenshots of possible new designs)* [FG1P2, male, age 83, T2D]

It was also felt, by 2 (25%) participants, that "jargon" should be avoided as much as possible and that any medical terms that are included should be accompanied by additional explanations:

> *So I think alongside each complication, you could have an explanation of what that is…It should say what it specifically is—what it affects, I think it should give a bigger explanation.* [FG2P2, male, age 66, T1D]

> *It needs to assume a low level of jargon knowledge.* [FG2P1, male, age 58, T1D]

Development of the Prototype Site

User perceptions and preferences collected during focus group discussions were considered and discussed by the researcher (CS), several diabetes clinicians (NC and DW), and the lead platform developer (DB). Those that were felt by all members of the team to be useful and feasible were incorporated into the prototype site. In this way, user insights directly influenced prototype site development. For example, participants' desire for clear visual cues and easy navigation was addressed by giving the homepage a "clean" look, with colored blocks (containing representative icons) for each test result/lifestyle item. When clicked, the blocks took the user to further details about each item. In comparison to the current MDMW site that participants were used to, fewer clicks were required to reach frequently sought information (eg, HbA_{1c} history graphs and eye and foot screening results).

Participants' desire for positive language was also heeded. An example in the risk prediction tool was the addition of a heading: "Let's reduce your risk." In response to participants' demand for plain language (and for additional explanations where medical terms were necessary), diabetes-related complications in the risk prediction tool were condensed into a single (combined) risk of developing complications, with an information icon alongside to explain what kinds of complications this pertains to. The positive reaction to the risk prediction tool as a means of fostering suitable goals was harnessed by adding the option (via a button) to link outputs from this page to the goal-setting area.

Think Aloud Sessions

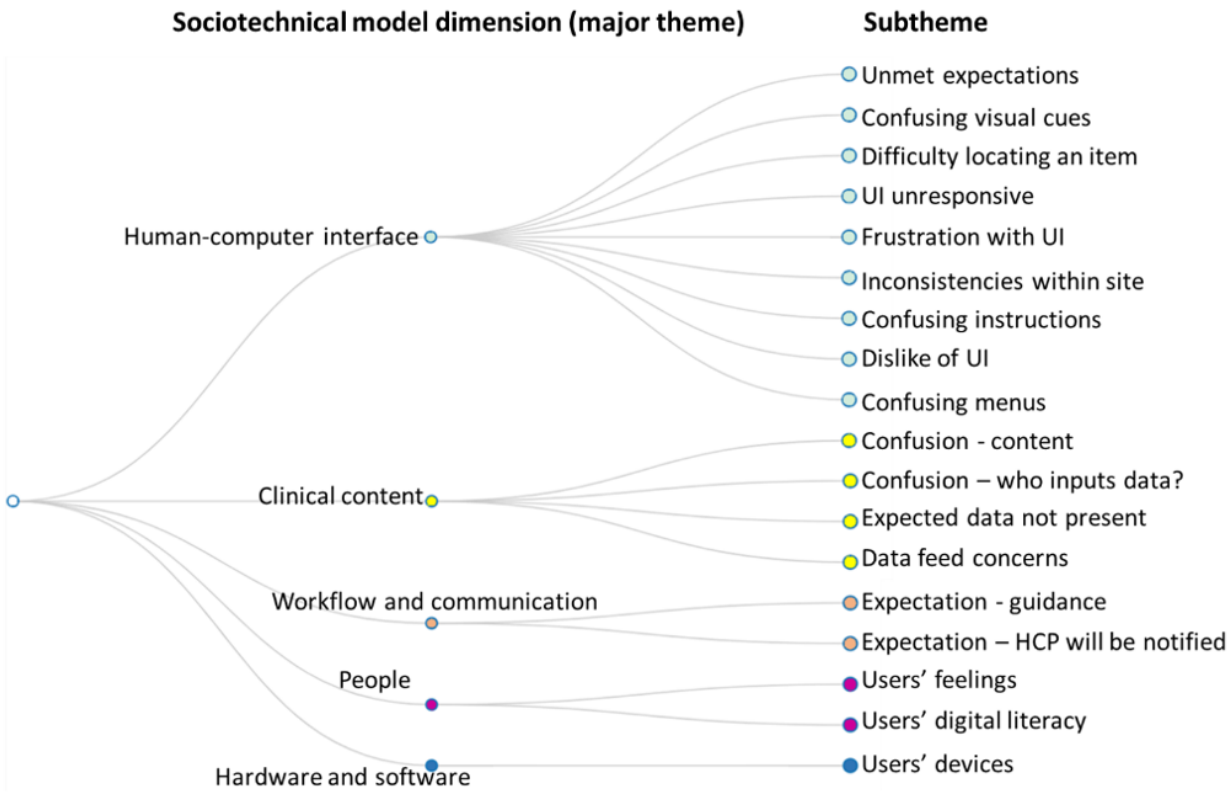
Think Aloud Participant Demographics

Seven participants took part in think aloud sessions with the prototype site. Of these, most were female (4/7, 57%), and most had T1D (5/7, 71%). The mean age was 52 years (range 35-62 years), and the mean time since diagnosis was 23 years (range 4-37 years). Sessions had an average length of 41 minutes.

Thematic Analysis

The 126 pertinent text fragments were mapped onto the 5 relevant sociotechnical model dimensions, which formed the basis of the major (deductive) themes of the analysis. Through an inductive approach, subthemes were generated from the data within each of the 5 dimensions. Some fragments were coded into more than one thematic category, resulting in 136 fragments included in the analysis. Categories from the thematic analysis are summarized in Figure 3.

Figure 3. Thematic categories of think aloud transcript data. The major (deductive) themes were the five relevant dimensions of the sociotechnical model. Subthemes were (inductive) themes within each dimension that were generated from the data. HCP: health care provider; UI: user interface.



Usability issues identified from think aloud user testing are summarized below and are discussed in relation to each of the sociotechnical model dimensions.

Human-Computer Interface

The majority (79/136, 58%) of text fragments from think aloud transcripts mapped onto the human-computer interface dimension of the sociotechnical model. Within this dimension, 9 subthemes were identified (Figure 3). One of the most prevalent of these, with 17 (13%) text fragments, was "confusing visual cues." Examples include users not realizing that an item on a page was clickable and it not being obvious what the different areas of a page represent:

> *As I hover over it, it's 'coming up' which suggests it is clickable, but it's not that obvious. I wouldn't…it's not that obvious. (referring to an object's hover state)* [TA_P1, female, age 35, T1D]

> *One of my first thoughts was: it's not obvious. It needs to say 'current' and then 'future.' (referring to left-and right-hand sides of the risk prediction tool page)* [TA_P2, female, age 49, T1D]

Another common theme, 15 (11%) text fragments, within the human-computer interface dimension was "unmet expectations." An example (in the risk prediction tool) was the expectation by users to be able to select a value for lifestyle/clinical variables and type in a replacement value, rather than changing this using a slider:

> *It doesn't seem to highlight 'HbA$_{1c}$'. Now with the 'smoking' you had the wee bubble so you could change it. And you had a slider bar for 'activity'. But with 'HbA$_{1c}$, you've got nothing, so I don't know whether you click on it.* [TA_P4, female, age 49, T1D]

"Inconsistencies within the site" was also a recurring theme within the human-computer interface dimension, 6 (4%) text fragments. An example was that some blocks on the homepage altered when hovered over (a shadow appeared around the edges), indicating to the user that they were clickable. Other

blocks did not possess altered hover states, leading some users to assume these were not clickable.

Some items were unresponsive, leading to several users becoming frustrated. An example was a slider in which the value did not change as the user moved the slider button along and only refreshed once the button was released:

> *Right so when you're actually clicking on the slider, although you can move it, it's not moving with the number.* [TA_P2, female, age 49, T1D]
>
> *It doesn't change as you move the slider. No…It's very, very fiddly.* [TA_P4, female, age 49, T1D]

Clinical Content

Six text fragments (4%) referred to "confusion over who inputs data" (ie, whether it is the user or their health care team who should do this). Within the "managing your condition" page (a checklist of diabetes tests, checks, and services that all people with diabetes should receive), some items (eg, "have your legs and feet checked" and "have your blood pressure measured") are automatically updated by the system once the user has attended a screening appointment, while others (eg, "receive high-quality care if admitted to hospital" and "get emotional and psychological support") must be updated by the user.

> *I thought that someone at the clinic would be doing that for you…it would be quite nice to have something to say that can be completed by the patient.* [TA_P5, female, age 51, T1D]

Within "clinical content," there was also a subtheme of "expected data not being present." For example, 2 (29%) participants expected access to retinal scan images via the retinal status page.

Workflow and Communication

Three of the participants (43%) wrongly expected that their HCP would be notified when they entered data into the system. For example, within the "managing your condition" page, the user can tick "received" or "not received" alongside a checklist item such as "get emotional and psychological support." Some users incorrectly assumed that ticking "not received" would alert their GP that they needed further support. This is not the case, as the checklist is intended to serve only as a reminder for the user:

> *So basically it gives the patient an opportunity to ask for something which they think could be useful to them.* [TA_P3, male, age 62, T2D]

People

Four participants (57%) pointed out that users with limited digital skills might find it challenging to interact with some aspects of the DHI. This was mentioned in relation to using the risk prediction tool (which involves the user interacting with several sliders and toggles and interpreting risk prediction outputs), as well as in relation to navigating from the homepage to other pages, as in the following example:

> *Having green boxes and red boxes is really good. I'm all for colors and highlighting things. Especially for people who are not as tech-savvy, as I am not.* [TA_P4, female, age 49, T1D]

Sentiments that the risk prediction tool could potentially cause anxiety, which had been expressed during focus groups, were reiterated by 4 (57%) think aloud participants:

> *It might put a lot of pressure on some people because it's not as easy to change your HbA_{1c} as just moving a slider.* [TA_P5, female, age 51, T1D]
>
> *It's important it's not rolled out to some patients who have a high risk and who have little or no opportunity to address that.* [TA_P6, male, age 58, T1D]

Two participants (29%) commented that additional information that appears beneath page titles (eg, the HbA_{1c} page and weight and BMI page) explaining these medical/lifestyle terms were superfluous and could make users feel "patronized":

> *The banner headline at the top. It's very patronizing. Stating the bleeding obvious, you know? For anyone who's been diagnosed for a while and has a bit of understanding about the condition.* [TA_P6, male, age 58, T1D]

Hardware and Software

Three text fragments (2%) referred to users' experience of the site being dependent on the device they use. One example is a list of screening results that appeared at the bottom of the retinal status page, which one participant felt might be missed by someone accessing the site using a mobile phone:

> *Unless you know to scroll down you might miss that bit about retinopathy. I suppose if you were looking on a phone, as well, you might not see that.* [TA_P5; female, age 51, T1D]

Development of the Final Site

User insights gained from focus groups and think aloud workshops with the prototype site were discussed among the researcher, clinicians, and developer (CS, NC, DW, and DB), and those that were feasible were used to inform the development of the final (live) MDIQ-augmented MDMW site. The resulting changes included the use of more positive and encouraging language throughout the site and mitigating users' anxieties. An example of the latter is a line of text added to the risk prediction tool suggesting that the user talks to their health care team if they are concerned about anything they have seen (Table 1). A user's risk of developing complications will be presented as a single "combined risk," with a breakdown of what those complications are (for users who wish to find out more). A simple explanation of what probability means will also be explained in a text box via an information icon (Table 1). A video will be added alongside the risk prediction tool, explaining how to use the tool, defining the terms used within the tool, and what the outputs mean. This will assist users with low digital and health literacy and will satisfy the need for plain language. The risk prediction tool will be displayed only after a new user's third visit to the MDMW site, allowing newly diagnosed users to become familiar with the site (and with having diabetes) before being presented with their complications risk forecast.

In the final site, if a user saves the lifestyle settings in the risk prediction tool as goals, these will be saved to the goal-setting area of the site. Within the goal-setting area, such goals will appear alongside links to resources on setting achievable goals, as well as links to targeted education resources (eg, resources on smoking cessation for those who set a goal to "give up smoking").

Several examples of unclear visual cues have been addressed in the final site. Within the risk prediction tool, the representation of "current risk" on the left-hand side of the risk prediction tool and potential "future risk" on the right-hand side, which had been unclear to several think aloud participants, has been made explicit (left side has been made smaller, and there are now clear labels above each side). Signposting to sliders within the risk prediction tool has also been improved (Table 1). Sliders have been made more consistently styled and more responsive (changing in real time as the slider button is dragged) in comparison with those in the prototype site.

Table 1. Examples of how insights gained during focus groups and think aloud workshops informed the development of the final MDIQ[a]-augmented MDMW[b] site.

Insight from focus groups/think aloud workshops	Resulting change in live MDIQ-augmented site
User being shown their risk of developing diabetes-related complications could cause anxiety.	Addition of text box to the risk prediction tool: 
Plain language needed, explain terms used	 Information icon text, when clicked: 
Clearer instructions needed	For example, addition of text directly above the sliders on the risk prediction tool: 

[a]MDIQ: MyDiabetesIQ.

[b]MDMW: MyDiabetes My Way.

Discussion

Principal Findings

This study has demonstrated that end-user involvement at each iterative stage of the design of a DHI can help to prioritize users' requirements throughout the design process. This novel version of an existing DHI for diabetes (MDMW) will be underpinned by an ML engine (MDIQ). Before developing a prototype interface, developers were made aware of user perceptions and preferences collected during initial focus groups. The resulting prototype was then user tested through the think aloud method, leading to the development of a new live MDIQ-augmented MDMW site, soon to be released. This work paves the way for a future iteration of this user-centered design process, which will entail large-scale real-world testing of the live site.

The sociotechnical model used here as a framework for the thematic analysis of think aloud data [22] needed to be adapted, with 3 of the model dimensions ("internal organizational policies, procedures, culture and environment"; "external rules, regulations, and pressures"; "system measurement and monitoring") not being relevant in the context of this patient-facing DHI. These dimensions are more applicable in clinician-facing health information technology used in formal health care settings, in which the model was originally conceived [22]. Think aloud data were therefore mapped onto the remaining 5 dimensions of the model ("human-computer interface," "clinical content," "workflow and communication," "people," and "hardware and software"), and subthemes were generated from the data.

It is useful to consider our findings in relation to each of the dimensions of the sociotechnical model individually. However, there is much interplay between dimensions. For example, many of the subthemes described within the human-computer interactions dimension (eg, "unclear visual cues" and "unclear instructions") would be compounded for a user who has limited digital skills (people dimension). Confusion over who inputs data (clinical dimension) could be avoided by addressing issues around human-computer interactions, including visual cues and instructions. Understanding the connections between the model dimensions can help developers find solutions that don't focus on one dimension while ignoring the impact of a given solution on others (ie, unintended consequences) [37]. This can additionally help developers to understand the wider sociotechnical structures already in place that could aid or constrain adoption of the technology under development [23] (eg, realizing that cues might be obvious to users who are used to interacting with online platforms but may not be obvious to less technically literate users). Awareness of these interactions between dimensions can help to make the site more usable for all users.

This study has demonstrated that viewing usability problems through a sociotechnical lens, and considering links between

sociotechnical dimensions, can help foster the development of a more acceptable DHI. Findings from prototype usability testing are being used to inform the development of the final site, such as when an unclear visual cue was identified (human-computer interactions dimension), it was amended to address concerns regarding digital literacy (people dimension). Similarly, users expressed some confusion regarding which data items were updated automatically and which items were to be updated by the user (clinical dimension). There was also confusion over which self-entered data items were fed back to the health care team (workflow and communication dimension). The user interface will be amended to clarify both and to mitigate against clinical risk.

This study found that participants were more receptive to users setting goals based on the outputs generated by the risk prediction tool compared to straightforward goal setting, where a user simply sets a goal for themselves ("potential to empower users" theme). Some participants were concerned about users setting "unrealistic" goals, thereby "setting themselves up for failure," or setting "unsafe" goals (people dimension). In response, the DHI now incorporates a link to the goal-setting area within the risk prediction tool and has been amended to highlight the need to set achievable goals, with the aim to provide a "scaffold" to foster healthy goals whilst aligning these goals with the user's accepted level of risk of diabetes complications.

Participant desires may not always be actionable, particularly when they rely on third parties whose working practices cannot be dictated by the DHI developers; for example, validation and approval of all user-set goals by clinicians were deemed desirable by some participants but would require clinicians to have extra time and flexibility in their working practices to facilitate this (workflow and communication dimension).

The possibility of the system offering ongoing feedback on progress toward user-set goals was suggested to focus group participants. The platform will continue to be developed with the potential to incorporate goal-setting notifications and alerts via email and mobile devices in the future. Any such developments will need to address participants' concerns regarding the potential to induce anxiety or feelings of failure in the event of a goal not being achieved.

During focus group discussions, most participants expressed a need for "plain language" (avoidance of medical jargon), as well as the addition of explanations alongside medical terms. This informed the decision to condense the initial information given to users, with additional information accessed via an information icon. There were, however, contrasting views around annotating items with further explanations, with some users feeling "patronized." PWD are a diverse group, and preferences will differ amongst individuals. For example, the needs of someone with longstanding diabetes will likely differ from someone who is newly diagnosed (ie, participants identified this latter group as having a lot to gain from the platform). MDMW already encompasses some degree of tailoring (eg, diabetes type, cholesterol, and blood pressure). Further tailoring (eg, user preference, diabetes duration, etc) is technically possible, although there is the potential that in doing so, the platform may become overly complicated.

Limitations of the Study

Attempts to recruit a more diverse group of participants via purposive sampling were not realized owing to poor response rates to the initial invite. The number of participants was relatively small and skewed towards older people with T1D who have had diabetes for many years. However, due to the rich data it delivers, the think aloud method requires only small numbers of participants, a suggested sample size of 5 to 8 participants, to uncover a high proportion of usability problems [9]. In addition, the participants were all considered "expert patients" whose experience provided valuable insights.

Social distancing necessitated by the COVID-19 pandemic resulted in the use of remote online user testing. This approach has known disadvantages, such as moderators needing to deal with unexpected technical issues during sessions and participants potentially facing a cognitively demanding environment (ie, navigating a video conferencing tool as well as testing the online intervention in question) [38]. Efforts were made to minimize these issues by giving clear instructions prior to sessions and by reassuring participants during sessions. In addition, remote online user testing provides several potential benefits: participants can be observed using the intervention in a more authentic context compared with sitting in a lab with researchers, and participants have more control over the session than they would in a lab setting (eg, muting their audio) [38].

Conclusions

This study demonstrates that acting on the input of end users at each stage of development can help to build more acceptable DHIs, aligning features to users' needs and avoiding unintended consequences that might cause disengagement. Good control of diabetes already relies heavily on self-management. The disruption of clinical services secondary to the COVID-19 pandemic only served to accentuate the importance of diabetes self-management. Digital tools such as MDMW (and particularly the MDIQ-enhanced version now being developed) can offer personalized support for self-management, alongside access to patients' electronic health records in a user-friendly environment. These features can facilitate self-management, thereby reducing users' risk of developing diabetes-related complications (with potential significant reductions in health care costs). This study serves as an exemplar of user-centered design that will ensure that MDMW is relevant and usable for people with diabetes.

Acknowledgments

We would like to thank the My Diabetes My Way (MDMW) users who took part in this study. Thanks also to Louise McIver, who assisted in moderating think aloud sessions.

Conflicts of Interest

SGC and DJW are cofounders and shareholders of MyWay Digital Health. DB is an employee of MyWay Digital Health.

Multimedia Appendix 1
Focus group schedule.
[DOCX File , 14 KB - humanfactors_v9i1e29973_app1.docx]

Multimedia Appendix 2
Tasks evaluated in the think aloud sessions.
[DOCX File , 15 KB - humanfactors_v9i1e29973_app2.docx]

References

1. International Diabetes Federation. IDF DIABETES ATLAS, Ninth Edition. URL: https://diabetesatlas.org/en/ [accessed 2021-06-20]
2. Garg S. The "Economic Health" of U.S. Health Care and Role of Middlemen. Diabetes Technology & Therapeutics 2020 Feb 01;22(2):136-141. [doi: 10.1089/dia.2019.0395] [Medline: 31674806]
3. Alberti K, Zimmet P. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO Consultation. Diabet Med 1998 Jul;15(7):539-553. [doi: 10.1002/(sici)1096-9136(199807)15:7<539::aid-dia668>3.0.co;2-s] [Medline: 9686693]
4. Ben-Zeev D, Kaiser SM, Brenner CJ, Begale M, Duffecy J, Mohr DC. Development and usability testing of FOCUS: A smartphone system for self-management of schizophrenia. Psychiatric Rehabilitation Journal 2013 Dec;36(4):289-296. [doi: 10.1037/prj0000019] [Medline: 24015913]
5. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth App for the Self-management of Adolescent Type 1 Diabetes: A Pilot Study. J Med Internet Res 2012 May 08;14(3):e70. [doi: 10.2196/jmir.2058] [Medline: 22564332]
6. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). Journal of Biomedical Informatics 2016 Apr;60:243-251. [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
7. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication with Computer Technology. Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 2013.
8. Henshall C, Davey Z, Jacelon C, Martin C. A usability study to test the effectiveness, efficiency and simplicity of a newly developed Internet-based Exercise-focused Health App for Lung cancer survivors (iEXHALE): Protocol paper. Health Informatics J 2019 Oct 21;26(2):1431-1442. [doi: 10.1177/1460458219882268] [Medline: 31631739]
9. Georgsson M, Staggers N. An evaluation of patients' experienced usability of a diabetes mHealth system using a multi-method approach. J Biomed Inform 2016 Feb;59:115-129 [FREE Full text] [doi: 10.1016/j.jbi.2015.11.008] [Medline: 26639894]
10. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. International Journal of Medical Informatics 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
11. NHS England. Five Year Forward View. 2014. URL: http//Www.England.Nhs.Uk/Wpcontent/Uploads/2014/10/5yfv-Web.Pdf [accessed 2021-06-20]
12. Pelizzola D. Diabetes assistance before, during and after Covid-19 in Ferrara, Italy. International Journal of Care Coordination 2020 Sep 11;23(2-3):61-64. [doi: 10.1177/2053434520954614]
13. Fuchs J, Hovorka R. COVID-19 and Diabetes: Could Diabetes Technology Research Help Pave the Way for Remote Healthcare? J Diabetes Sci Technol 2020 May 30;14(4):735-736. [doi: 10.1177/1932296820929714] [Medline: 32475169]
14. Fico G, Martinez-Millana A, Leuteritz J, Fioravanti A, Beltrán-Jaunsarás ME, Traver V, et al. User Centered Design to Improve Information Exchange in Diabetes Care Through eHealth : Results from a Small Scale Exploratory Study. J Med Syst 2019 Nov 18;44(1):2. [doi: 10.1007/s10916-019-1472-5] [Medline: 31741069]
15. Cafazzo JA. A Digital-First Model of Diabetes Care. Diabetes Technology & Therapeutics 2019 Jun;21(S2):S2-52-S2-58. [doi: 10.1089/dia.2019.0058] [Medline: 31169428]
16. Fu HNC, Rizvi RF, Wyman JF, Adam TJ. Usability Evaluation of Four Top-Rated Commercially Available Diabetes Apps for Adults With Type 2 Diabetes. CIN Comput Informatics, Nurs 2020;38(6):274-280. [doi: 10.1097/cin.0000000000000596] [Medline: 31904594]
17. Eyles H, Jull A, Dobson R, Firestone R, Whittaker R, Te Morenga L, et al. Co-design of mHealth Delivered Interventions: A Systematic Review to Assess Key Methods and Processes. Curr Nutr Rep 2016 Jul 4;5(3):160-167. [doi: 10.1007/s13668-016-0165-7]
18. Marent B, Henwood F, Darking M. Development of an mHealth platform for HIV Care: Gathering User Perspectives Through Co-Design Workshops and Interviews. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e184. [doi: 10.2196/mhealth.9856] [Medline: 30339132]

19. World Health Organization. Monitoring and evaluating digital health interventions: a practical guide to conducting research and assessment. 2016. URL: https://apps.who.int/iris/handle/10665/252183 [accessed 2020-06-20]
20. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. Journal of Biomedical Informatics 2013 Dec;46(6):1080-1087. [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
21. Eshet E, Bouwman H. Context of Use: The Final Frontier in the Practice of User-Centered Design? Interact Comput 2016 Oct 13:368-390. [doi: 10.1093/iwc/iww030]
22. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Quality and Safety in Health Care 2010 Oct 18;19(Suppl 3):i68-i74. [doi: 10.1136/qshc.2010.042085] [Medline: 20959322]
23. Sheehan B, Nigrovic LE, Dayan PS, Kuppermann N, Ballard DW, Alessandrini E, et al. Informing the design of clinical decision support services for evaluation of children with minor blunt head trauma in the emergency department: A sociotechnical analysis. Journal of Biomedical Informatics 2013 Oct;46(5):905-913. [doi: 10.1016/j.jbi.2013.07.005] [Medline: 23892207]
24. Baldwin J, Singh H, Sittig D, Giardina T. Patient portals and health apps: Pitfalls, promises, and what one might learn from the other. In: Healthcare. Amsterdam: Elsevier; 2017:81-85.
25. Giardina T, Baldwin J, Nystrom D, Sittig D, Singh H. Patient perceptions of receiving test results via online portals: a mixed-methods study. J Am Med Informatics Assoc 2018;25(4):440-446. [doi: 10.1093/jamia/ocx140] [Medline: 29240899]
26. Hägglund M, Scandurra I. A socio-technical analysis of patient accessible electronic health records. Stud. Health Technol. Inform 2017;244:3-7. [Medline: 29039366]
27. Carayon P, Wetterneck TB, Rivera-Rodriguez AJ, Hundt AS, Hoonakker P, Holden R, et al. Human factors systems approach to healthcare quality and patient safety. Applied Ergonomics 2014 Jan;45(1):14-25. [doi: 10.1016/j.apergo.2013.04.023] [Medline: 23845724]
28. Murphy K. Machine Learning: A Probabilistic Perspective. Cambridge, Massachusetts: MIT press; 2012:A.
29. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. The Lancet 2019 Apr;393(10181):1577-1579. [doi: 10.1016/s0140-6736(19)30037-6] [Medline: 31007185]
30. Zipkin DA, Umscheid CA, Keating NL, Allen E, Aung K, Beyth R, et al. Evidence-Based Risk Communication. Ann Intern Med 2014 Aug 19;161(4):270. [doi: 10.7326/m14-0295] [Medline: 25133362]
31. Wake DJ, He J, Czesak AM, Mughal F, Cunningham SG. MyDiabetesMyWay. J Diabetes Sci Technol 2016 Jul 09;10(5):1050-1058. [doi: 10.1177/1932296816648168] [Medline: 27162192]
32. Cunningham SG, Brillante M, Allardice B, Conway N, McAlpine RR, Wake DJ. My Diabetes My Way: supporting online diabetes self-management: progress and analysis from 2016. BioMed Eng OnLine 2019 Feb 8;18(1):1-11. [doi: 10.1186/s12938-019-0635-4] [Medline: 30736798]
33. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. Albar SA, Carter MC, Alwan NA, Evans CEL, Cade JE. Formative evaluation of the usability and acceptability of myfood24 among adolescents: a UK online dietary assessments tool. BMC Nutr 2015 Oct 23;1(1):1-11. [doi: 10.1186/s40795-015-0016-8]
35. Ericsson K, Simon HA. Protocol Analysis: Verbal Reports as Data. Cambridge, Massachusetts: The MIT Press; 1984.
36. Gilhooly K, Fioratou E, Anthony S, Wynn V. Divergent thinkingtrategies and executive involvement in generating novel uses for familiar objects. Br J Psychol 2007;98(4):611-625. [doi: 10.1111/j.2044-8295.2007.tb00467.x] [Medline: 17535464]
37. Ash JS, Sittig DF, Poon EG, Guappone K, Campbell E, Dykstra RH. The Extent and Importance of Unintended Consequences Related to Computerized Provider Order Entry. Journal of the American Medical Informatics Association 2007 Jul 01;14(4):415-423. [doi: 10.1197/jamia.m2373] [Medline: 17460127]
38. Wozney LM, Baxter P, Fast H, Cleghorn L, Hundert AS, Newton AS. Sociotechnical Human Factors Involved in Remote Online Usability Testing of Two eHealth Interventions. JMIR Human Factors 2016 Feb 03;3(1):1-12. [doi: 10.2196/humanfactors.4602] [Medline: 27026291]

Abbreviations

AI: artificial intelligence
DHI: digital health interventions
GP: general practitioner
HbA_{1c}: glycated hemoglobin
HCP: health care professional
MDIQ: MyDiabetesIQ
MDMW: My Diabetes My Way
ML: machine learning
NHS: National Health Service

PPI: patient and public involvement
PWD: people with diabetes
T1D: type 1 diabetes
T2D: type 2 diabetes

Edited by A Kushniruk; submitted 27.04.21; peer-reviewed by I Gabashvili, D Chrimes; comments to author 02.10.21; revised version received 03.11.21; accepted 06.11.21; published 08.02.22.

Please cite as:
Shields C, Cunningham SG, Wake DJ, Fioratou E, Brodie D, Philip S, Conway NT
User-Centered Design of A Novel Risk Prediction Behavior Change Tool Augmented With an Artificial Intelligence Engine (MyDiabetesIQ): A Sociotechnical Systems Approach
JMIR Hum Factors 2022;9(1):e29973
URL: https://humanfactors.jmir.org/2022/1/e29973
doi:10.2196/29973
PMID:35133280

Original Paper

Patients' and Providers' Perspectives on and Needs of Telemonitoring to Support Clinical Management and Self-care of People at High Risk for Preeclampsia: Qualitative Study

Maria Aquino[1,2], BScN, MHI; Janessa Griffith[3,4], BA, MSc, PhD; Tessy Vattaparambil[1,2], MHI; Sarah Munce[5,6], PhD; Michelle Hladunewich[7], MD, MSc; Emily Seto[1,2], PEng, PhD

[1]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[2]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

[3]Institute of Medical Science, University of Toronto, Toronto, ON, Canada

[4]Women's College Hospital Institute for Health System Solutions and Virtual Care, Women's College Hospital, Toronto, ON, Canada

[5]Department of Occupational Sciences and Occupational Therapy, University of Toronto, Toronto, ON, Canada

[6]KITE-Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

[7]Division of Nephrology, Department of Medicine, Sunnybrook Health Sciences Centre, Toronto, ON, Canada

Corresponding Author:
Maria Aquino, BScN, MHI
Institute of Health Policy, Management and Evaluation
Dalla Lana School of Public Health
University of Toronto
155 College St 4th Floor
Toronto, ON, M5T 3M6
Canada
Phone: 1 (416) 978 4326
Email: maria.aquino@mail.utoronto.ca

Abstract

Background: Preeclampsia is one of the leading causes of maternal mortality worldwide, with a global prevalence at 2%-8% of pregnancies. Patients at high risk for preeclampsia (PHRPE) have an increased risk of complications, such as fetal growth restriction, preterm delivery, abnormal clotting, and liver and kidney disease. Telemonitoring for PHRPE may allow for timelier diagnosis and enhanced management, which may improve maternal and perinatal outcomes.

Objective: The objective of this study is to determine the perceptions and needs of PHRPE and their health care providers with respect to telemonitoring through semistructured interviews with both groups. This study explored (1) what the needs and challenges of monitoring PHRPE are during pregnancy and in the postpartum period and (2) what features are required in a telemonitoring program to support self-care and clinical management of PHRPE.

Methods: This study used a qualitative descriptive approach, and thematic analysis was conducted. PHRPE and health care providers from a high-risk obstetrical clinic in a large academic hospital in Toronto, Canada, were asked to participate in individual semistructured interviews. Two researchers jointly developed a coding framework and separately coded each interview to ensure that the interviews were double-coded. The software program NVivo version 12 was used to help organize the codes.

Results: In total, 7 PHRPE and 5 health care providers, which included a nurse practitioner and physicians, participated in the semistructured interviews. Using thematic analysis, perceptions on the benefits, barriers, and desired features were determined. Perceived benefits of telemonitoring for PHRPE included close monitoring of home blood pressure (BP) measurements and appropriate interventions for abnormal BP readings; the development of a tailored telemonitoring system for pregnant patients; and facilitation of self-management. Perceived barriers to telemonitoring for PHRPE included financial and personal barriers, as well as the potential for increased clinician workload. Desired features of a secure platform for PHRPE included the facilitation of self-management for patients and decision making for clinicians, as well as the inclusion of evidence-based action prompts.

Conclusions: The perceptions of patients and providers on the use of telemonitoring for PHRPE support the need for a telemonitoring program for the management of PHRPE. Recommendations from this study include the specific features of a

telemonitoring program for PHRPE, as well as the use of frameworks and design processes in the design and implementation of a telemonitoring program for PHRPE.

(JMIR Hum Factors 2022;9(1):e32545) doi:10.2196/32545

KEYWORDS

high-risk pregnancy; blood pressure; preeclampsia; telemonitoring; home monitoring; mHealth

Introduction

Preeclampsia is a progressive multisystem syndrome that involves new-onset hypertension and end-organ dysfunction or proteinuria in the second half of pregnancy or postpartum [1]. The prevalence of preeclampsia is 2%-8% worldwide [2]. The prevalence is higher in first-time pregnancies, in people of advanced maternal age, in people who have had preeclampsia in the past, and in those with preexisting medical conditions (eg, chronic hypertension, kidney disease, diabetes mellitus) [3]. People who are diagnosed with preeclampsia are at an increased risk of both obstetrical complications, such as fetal growth restriction and preterm delivery [4], as well as medical complications, such as abnormal clotting, peripheral and pulmonary edema, and liver and kidney dysfunction [5]. Additionally, preeclampsia increases the long-term risk of developing hypertension, ischemic heart disease, stroke, venous thromboembolism, kidney failure, and death [6-8]. Hypertensive disorders of pregnancy, including chronic hypertension, gestational hypertension, and preeclampsia [8], are a leading cause of maternal mortality worldwide [7]. Identifying patients at high risk for preeclampsia (PHRPE) may improve the probability of a timely diagnosis, which may enhance maternal and perinatal outcomes [9].

Home blood pressure (BP) monitoring may play a role as an adjunct to managing PHRPE for developing preeclampsia, because it allows for more frequent BP readings and may potentially lead to timelier detection of preeclampsia [10]. Home BP monitoring may also be beneficial in assessing the BP control of pregnant people with chronic hypertension [10]. The 2013 American College of Obstetricians and Gynecologists Guidelines recommend home BP monitoring for pregnant people with gestational hypertension, chronic hypertension, and uncontrolled BP [1].

Digital health interventions (DHIs), such as telemonitoring, can provide a personalized approach to the outpatient monitoring of PHRPE [11,12]. Specifically, telemonitoring enables health care providers to have access to real-time information about their patient's home BP readings for clinical decision support, such as medication changes or delivery of the baby. Telemonitoring can also enable targeted and timely self-management instructions for PHRPE based on their current and trending BP readings. In our scoping review, published in 2020, we identified only 20 studies exploring telemonitoring of patients at high risk for hypertensive disorders of pregnancy [13]. These papers described telemonitoring interventions in countries such as the United Kingdom, the United States, and Belgium, but there were no Canadian studies published on telemonitoring interventions for patients at high risk for hypertensive disorders of pregnancy at the time [13]. Although we determined that telemonitoring could provide benefits for managing patients at high risk for hypertensive disorders, more research is needed to understand how to develop and implement such programs, as well as to prove its safety and effectiveness [13]. To the best of our knowledge, there has been no recent Canadian study published on the topic since our scoping review. Recently, 3 studies from the United Kingdom have explored the perspectives of women [14] and health care providers [15] as well as a design and implementation approach [16] for BP self-monitoring.

The overall objective of this study is to determine the perceptions and needs of PHRPE and their health care providers with respect to telemonitoring. There were 2 specific research questions for this study: (1) What are the needs and challenges of monitoring PHRPE during pregnancy or in the postpartum period? (2) What are the features required in a telemonitoring program to support self-care and clinical management of PHRPE?

Methods

Study Design

This study adopted a qualitative descriptive approach [17]. PHRPE and health care providers involved in their care from a high-risk obstetrical clinic at a large academic hospital in Toronto, Canada, participated in individual semistructured interviews. The study was approved by the research ethics boards of the University Health Network (#18-5535) and the Sunnybrook Health Sciences Centre (#295-2018). This research followed the guidance of the Consolidated Criteria for Reporting Qualitative Research [18].

Participants and Recruitment

This study used purposive sampling with PHRPE and health care providers. PHRPE were eligible for the study if they were identified as being at high risk for preeclampsia, aged 18 years or older, able to communicate in English, pregnant or up to 6 weeks postpartum, and currently self-monitoring home BP measurements. Eligible patients were identified and were informed about the study by their physician during a clinic visit and were asked whether they would be willing to speak to the research coordinator (author MA) to participate in the study. The researchers had no prior relationship with the participants. The patients were also provided with a letter of invitation with further information that stated that their health providers would not know the identity of the study participants. Health care providers associated with the high-risk obstetrical clinic who were involved in the care of PHRPE were also invited to participate in a semistructured interview. They were invited to participate via email by the clinical site lead and could have various clinical roles, including general practitioners,

obstetricians, nurse practitioners, nurses, and pharmacists. PHRPE and health care providers interested in participating were contacted by the research coordinator, who provided the details of the study and answered any questions of the potential participants prior to obtaining written consent. Recruitment continued until the data had reached saturation, which was anticipated to be between 5 and 15 participants in each of the 2 groups.

Data Collection and Analysis

The interviews were conducted in person by the research coordinator (MA) in a private room at the clinic. The research coordinator had experience interviewing patient participants and had graduate-level training in research, as well as a background in nursing for the adult cardiac population. Each interview lasted between approximately 20 and 30 minutes. The interview guides are provided in Multimedia Appendix 1. The interviews were digitally recorded and transcribed verbatim by a professional transcriptionist for analysis. Repeat interviews did not occur.

Thematic analysis, as described by Braun and Clarke [19], was conducted. Two researchers (authors MA and JG, both she/her) coded several transcripts together to develop a common coding framework. Both researchers had advanced-level training for coding qualitative interviews. After the framework was established, the researchers separately coded each interview to ensure each interview was double-coded. The researchers discussed the codes to resolve any discrepancies. During this discussion, the codes were grouped into initial themes. The codes and themes were then refined through discussion by the larger research team. The software program NVivo version 12 (QSR International, Doncaster, Victoria, Australia) was used during the analysis of the transcripts to help organize the codes.

Results

Participant Demographics

In total, 7 patients participated in the study and ranged from being 12 weeks pregnant to 2 months postpartum. Many patient participants were referred to this clinic due to chronic kidney disease that caused hypertension during pregnancy. All patient participants measured their BP at home with varying degrees of frequency. The most common frequency of BP self-measurement was 3 times per day, with 3 (43%) of 7 patient participants using this schedule. Most patient participants also used a DHI before pregnancy to monitor other health conditions, such as diabetes, activity, and fertility. Only 1 (14%) of 7 patient participants used a mobile app, called QardioArm, to monitor their BP during pregnancy. Further details on the characteristics of the patient participants are shown in Table 1.

Most staff in the clinic also participated in the study, including specialists, staff physicians, and a nurse practitioner, with 4 (80%) of 5 providers being female. None of the providers had previously recommended the use of telemonitoring tracking technology for their patients. There were 5 provider participants, of whom 4 (80%) were female and 1 (20%) was male. The provider participants comprised 4 (80%) physicians and 1 (20%) nurse practitioner. In general, the health care providers were experienced in the fields of maternal-fetal medicine as well as nephrology.

Table 1. Demographics of patient participants.

Patient	Reason for referral to clinic	Frequency of BP[a] self-measurements	Use of DHI[b] before pregnancy	Use of DHI to monitor BP during pregnancy
1	Chronic kidney disease	3 times per day	No	No
2	Adrenal tumor, hypertension	6-10 times per day	Yes, for diabetes	No
3	Polycystic kidneys, hypertension	2 times per day	Yes, fertility tracker	Yes, QardioArm
4	Hypertension	3 times per day	Yes, step tracker	No
5	Nephrotic syndrome	3 times per day	No	No
6	Kidney disease	With symptoms	Yes, step tracker	No
7	Hypertension	2 times per day	Yes, Carrot app	No

[a]BP: blood pressure.

[b]DHI: digital health intervention.

Themes From the Interviews

During the interviews, participants discussed how they envisioned a telemonitoring program could be used in the care of patients at the clinic. The following themes emerged within the categories of perceived benefits of, perceived barriers to, and desired features of telemonitoring for PHRPE.

Perceived Benefits of Telemonitoring for PHRPE

The perceived benefits were as follows:

- Telemonitoring PHRPE may provide close monitoring of home BP measurements and appropriate interventions for abnormal BP readings: Providers indicated that patients at risk for preeclampsia require close and frequent monitoring between clinic appointments. Some providers stated that telemonitoring could allow for remote observation of patients who have already been identified as having BP values that could change suddenly and unpredictably.

  I think it can help some patients, yeah, I mean, the[se] patients [are] definitely one [of] the highest [risk] groups who really have to watch blood pressure quite

> *carefully. Yeah, so this wouldn't be, it wouldn't probably [be], something I would use with the majority of patients who are at risk but in terms of blood pressure in such an intensive manner. We're talking about patients who already either have chronic hypertension that needs to be managed and followed closely or patients who are already diagnosed with mild preeclampsia and are being managed as outpatients. That probably would be the target population.* [Provider 4]

Additionally, telemonitoring may allow for improved access to perinatal health services for patients who may not have easy access to these resources.

> *They're watched very carefully [in the] clinic, but there are women that are not watched carefully in other obstetrical population[s] . . . not everybody's got that kind of resource; so then, you have to think out [if it could] be even a more useful tool for women to self-identify and come to triage maybe in time before they get too sick.* [Provider 5]

- Telemonitoring PHRPE may provide tailored information and interventions for pregnant patients: Providers reported that currently available telemonitoring systems are not tailored for pregnant people, which causes provider hesitancy in recommending them to patients.

> *So there hasn't been [a telemonitoring system] that's been sort of really tailored right for our population . . . they haven't been sort of detailed enough for my pregnant women—so I think it could happen, but we don't have the right one yet.* [Provider 5]

When patients were asked how they currently manage their BP between clinic visits, some patients mentioned that they resorted to using internet search engines for information regarding symptoms they were experiencing. Patients found information that was not specific to preeclampsia. A BP telemonitoring system that is tailored specifically to patients at risk for preeclampsia was thought to be useful.

> *It's different criteria for different [patients]—like if you're diabetic . . . and you're a certain age or, you know, [there are] different criteria for different things . . . because what I find that even as a patient, you know, I find that even with a nursing background, sometimes I don't know when to go to emergency, you know. I'm pregnant now, and I'm bleeding . . . should I go to [the emergency department] or should I not', and I would imagine that's the same thing that other people who don't have a health care background may struggle with.* [Patient 3]

Patients expressed a desire to include contextual information related to their BP measurements (eg, medications taken, emotional and physical states, time of day) to provide their health care providers with more information about any potential influencing factors on the measurement.

> *I don't know if clinicians would want to know when you took a blood pressure medication, like did you take your blood pressure medication, like 'cause you don't [want] to assume that they did cause some people might not have. Or . . . if they want to have a comment section where you can write I've been feeling sick or I'm in a lot of pain or something like that. So people can write a comment next to the blood pressure results.* [Patient 7]

Providers noted that acquiring personal information about what patients are experiencing allows them to make better decisions on their care.

> *I think it's crucial keeping a diary in pregnancy and for preeclampsia . . . [It allows me] to adjust medications and [know] how severe things are.* [Provider 3]

- Telemonitoring for PHRPE may facilitate self-management in patients: Self-management in this context can include actions taken, as needed, based on symptoms and self-monitoring of BP. An example of self-management would be a pregnant patient going to the emergency department for higher-than-normal BP readings and accompanying symptoms, such as a headache. Both clinicians and patients indicated that a telemonitoring system that educates patients on self-management practices could be empowering and could build their self-confidence in managing their condition. The study participants expressed that this sense of empowerment could help create positive habits for patients, including regular self-monitoring of their BP and an awareness of actions to take when the results are outside of the normal range.

> *You have to also make sure you're empowering the patient to do self-management, so if they are just thinking somebody else is watching . . . then that's not actually self-management, so that's the key, is to find an app that works but still encourages [patients] to take a very active part in their health care.* [Provider 5]

> *I think [an app] empowers people to be more proactive about their health and what they should do and who they should reach out to. People may take their [blood pressure] readings or may not and . . . they don't know when to come in to see their family doctor or their specialists. It's also a waste of resources for emergency teams or the emergency department . . . when you go in unnecessarily; so I think it can be useful.* [Patient 3]

Perceived Barriers to Telemonitoring PHRPE

The perceived barriers were as follows:

- Financial and other personal barriers to telemonitoring: Clinicians and patients highlighted the need to address financial barriers (eg, cost of a home BP monitor) to enhance patient BP self-monitoring.

> *Some people would not have the financial means to get a home blood pressure device, and going to [the] pharmacy, it's time demanding, it's complicated.* [Provider 4]

> *And if the [telemonitoring is] covered under insurance, [because] they can get pricey . . . especially if the doctor said you need to [monitor your blood pressure at home], if [providers] can write [a] prescription and the private insurance companies . . . can cover up to 70% [of the cost] or something like that.* [Partner of patient 5]

A telemonitoring system would need to demonstrate an added benefit or value in the patient's care management, especially for those patients who currently manage their health without technology.

> *To be honest with you, it's, I guess, it's too much work, and I'm not really into, like, the whole technology thing.* [Patient 5]

> *I think the patient has to be convinced there is an issue, and I think sometimes the only way is if the patient is made to lead her condition and also to take charge. That's the only way I feel she would take it seriously. And I think making them record or making them understand more will help motivate them. So I think if that does exist, that app, or whatever program, I think it would actually really help, because we do ask patients to record [blood pressure measurements]. I ask them to record.* [Provider 3]

Finally, access to the appropriate equipment, such as validated BP cuffs, for pregnant patients may pose as a barrier to the telemonitoring of this group.

> *Not all blood pressure cuffs are created equally for pregnant women, so you'll have to validate your cuff first in a pregnant population because it may not be the same . . . that's why I have to [ask patients to] bring all their cuffs in to calibrate, so [patients] always bring their home [blood pressure] cuffs in, and we calibrate them [at the clinic].* [Provider 5]

- Telemonitoring PHRPE may increase clinician workload: Some providers were concerned that BP self-monitoring for patients using telemonitoring would increase clinician workload. Reasons given included being alerted to abnormal BP readings during off-clinic hours, as well as patients requiring timely assessments, follow-up, and recommendations based on BP readings.

> *I mean in theory [telemonitoring] would be good, but who's monitoring that is the problem . . . I think there's enough information for us to deal with, and the expectation, I would just be a little bit hesitant to say that the expectation is for the physician to monitor the blood pressure . . . in terms of data that the patients are inserting into their app, because then it becomes a safety thing, because I'm not checking my phone all the time and I don't think that should be the precedent.* [Provider 1]

Desired Features of a Telemonitoring Program

The desired features were as follows:

- Telemonitoring PHRPE should facilitate self-management for patients: Patients want to be able to easily transfer their BP results to a centralized location where it can be readily accessed by their provider. Both patients and providers commented on the ease they found using a diabetes monitoring system.

> *Because like me, I'm a diabetic, [and] my diabetes machine connects to my phone, so it's always synced there, so I see everything on my phone. And if I want to send it to my doctor, I just send a link to my doctor, and he can see every day what was my sugar.* [Patient 2]

> *Because when I think about what [patients] have to do, they have to create an email, like get an email and then write in all their readings. So if they could just have something that's like a blood sugar monitor that monitored them and kept track of it, and put it into an app, I think [patients] would be onboard with that because it would be easier for them than having to keep track of all their numbers, then transfer it over into an email.* [Provider 2]

- Telemonitoring of PHRPE should facilitate decision making for clinicians: Providers suggested displaying BP results, medications, and symptoms to enable easy identification of abnormal values, trends, and the appropriate intervention.

> *In a busy clinic, it's hard to sit and look at all the numbers, but imagine if patients plug a 2-week amount of blood pressure readings and then you get a graph. From that graph, you'll know the highest and the lowest [numbers], you'll know the trend, you'll know the timing of day, and so you can, first of all, if things are worse at the end of the pregnancy, it might tell you to deliver. If you're early on and showing you highest in the morning, then you might adjust nighttime [medication] dosing.* [Provider 3]

- Telemonitoring PHRPE should include evidence-based action prompts: Providers and patients highlighted the need for patient alerts and action prompts to be automatically generated from the patient data to ensure that patients seek medical attention, as needed.

> *I'm worried about [patients] throwing [blood pressure values] in [and] not knowing that, oh, like I'm waiting for somebody to get back to me, but meanwhile my blood pressures through the roof and I'm in danger, but because my app doesn't tell me to do anything, I stay home.* [Provider 1]

> *I think that [getting feedback is] really important because . . . it empowers people to be more proactive about their health and what they should do and who they should reach out to.* [Patient 3]

Providers described the need for a DHI that can incorporate evidence-based protocols and standards for PHRPE to facilitate their decision making. Specifically, a provider discussed an online risk calculator developed by the Fetal Medicine Foundation based in the United Kingdom where maternal risk factors and biomarkers are input to calculate a patient's risk for preeclampsia [20].

> *There is an online risk calculator that I personally use. It has been validated recently by a large study, published in [the] New England Medical Journal, so I actually use it to actually quantify the risk.* [Provider 4]

Discussion

Principal Findings

This paper presents the findings from a qualitative study aimed at determining the needs and challenges of monitoring PHRPE and the features required in a telemonitoring program to support the self-care and clinical management of PHRPE. Specifically, this study identifies participant perceptions on the benefits of, barriers to, and desired features of a telemonitoring program for PHRPE. Perceived benefits of telemonitoring for PHRPE included close monitoring of home BP measurements and appropriate interventions for abnormal BP readings; a tailored telemonitoring system for pregnant patients; and facilitation of self-management. Perceived barriers to telemonitoring for PHRPE included financial and personal barriers as well as the potential for increased clinician workload. Desired features of PHRPE included the facilitation of self-management for patients and decision making for clinicians, as well as the inclusion of evidence-based action prompts.

Patients and providers had similar opinions with respect to the features required in a telemonitoring program. For example, both patients and providers agreed that empowering patients is key to facilitating their self-management. Patients wanted feedback from a telemonitoring program so that they can feel empowered in their self-management of high BP during pregnancy. Patients looked to a telemonitoring program to educate themselves on what symptoms to identify, as well as the appropriate response when a symptom is identified. This viewpoint is similar to 2 recent studies published in the United Kingdom by Hinton et al [14]. One study described the acceptability and feasibility of self-monitoring BP during pregnancy from the women's perspective [14]. The participants in this feasibility study reported feeling reassured and empowered by self-monitoring their BP, especially if they had a history of hypertension or preeclampsia [14]. Study participants felt that BP self-monitoring made them more knowledgeable of the risks of hypertension and preeclampsia in pregnancy [14]. The second study interviewed 147 obstetricians, community and hospital midwives, pharmacists, and trainee doctors to gain their perspective on the use of home BP monitoring during pregnancy [15]. In this study, providers acknowledged the potential for self-monitoring of BP to empower women in their health.

In our study, patients and providers had differing views on the perceived barriers to a telemonitoring program with respect to patient motivation. Patients who were less inclined to use technology for their health indicated that it would be difficult to incorporate telemonitoring into their routine. Moreover, depending on a patient's situational context, such as the stage of pregnancy, incorporating telemonitoring into their life may become particularly challenging. For example, Hinton et al [14] noted that during the postpartum period, women found incorporating BP self-monitoring more challenging. Providers in our study noted that patient motivation is related to the patient's view of their disease and their lack of understanding related to its severity and impact on their health. A qualitative study by Davies et al [21] described primary care clinicians' views on telemonitoring and highlighted that a lack of patient motivation is a barrier to their use. However, our study, with its small sample size as a limiting factor, demonstrated a highly motivated group in that all 7 patient participants measured their BP at home.

Both patients and providers referred to the ease with which telemonitoring systems or mobile apps for patients with diabetes send blood sugar readings to their providers. In fact, a summary of the recent studies on telemonitoring projects for patients with diabetes indicated favorable results, including improved control of blood glucose levels and a significant reduction in hemoglobin A^{1c}, enhanced effects on comorbidities, increased quality of life for patients, and good uptake of the technology by patients [22].

A study published in 2019 by Band et al [16] incorporated theoretical modeling and a person-based approach to develop a logic model outlining the proposed mechanism of change for the Blood Pressure Self-Monitoring in Pregnancy (BUMP) program. This systematic approach to intervention development allowed for a deeper understanding and appreciation of the issues and experiences of pregnant patients with hypertensive disorders [16]. The logic model presents strategies, such as sending reminder prompts and cues to self-monitor and providing information about the benefits of self-monitoring, to address common barriers—barriers that were also reflected by the participants in our study to adopting self-monitoring during pregnancy [16].

Lastly, all providers were concerned about their clinical accountability with respect to a telemonitoring program for PHRPE. In fact, none of the providers had previously recommended the use of a telemonitoring program to their patients, stating the lack of availability of such telemonitoring systems and the lack of available guidelines for the telemonitoring of PHRPE. Similarly, Hinton et al [15] described the requirements identified by health care providers for telemonitoring programs to consider normal BP fluctuations throughout pregnancy and to ensure the accuracy of BP results in BP self-monitoring. Additionally, providers raised concerns about patients appropriately addressing abnormal BP [15]. These perspectives are also highlighted by a systematic review of the self-monitoring of BP in hypertension, which recommends the use of guidelines on how to interpret BP measurements taken at home, acceptable variances between home and clinic measurements, and how to direct patients on addressing concerning results [23].

Recommendations on Features of the Telemonitoring Program

Both patient and provider participant groups expressed interest in the potential benefits of a telemonitoring program for PHRPE. The following features represent recommendations that could be used in the design and development of future programs:

- Assess patients' technical literacy, motivation, and readiness to participate in a telemonitoring program preimplementation.
- Generate automatic alerts for patients when BP results are out of range, and instruct patients to either seek medical care or modify their self-care behaviors, depending on the BP value and trends, which would help address concerns related to clinician accountability.
- Develop alerts based on evidenced-based protocols and guidelines specific to PHRPE.
- Enable patients to add symptoms or other related information about their condition (eg, medications taken), in addition to BP readings, to better understand potential root causes for the symptoms.
- Send BP values automatically from the BP monitor to the telemonitoring system to avoid manual entry of data.
- Display patient data in a way that facilitates identification of trends and easy visualization of a patient's health status by the provider.

Recommendations on Frameworks for Program Implementation

Patient empowerment to facilitate self-management was a common theme discussed by both patient and provider participants. A framework that uses an empowerment-based approach to developing digital intervention tools for self-management has been described by Alpay et al [24]. This framework includes 6 components of patient empowerment (ie, communication, education and health literacy, information, self-care/support, decision aids, and contact with fellow patients) and may be beneficial in the implementation of a telemonitoring program for PHRPE [24]. This empowerment-based approach combined with the logic model, which was developed using a person-based approach proposed by Band et al [16], may provide direction on the design and implementation of a telemonitoring program for patients at high risk for hypertensive disorders of pregnancy.

Limitations

Although the sample size was small and homogenous, with 7 patients and 5 health care providers being interviewed from a single health care institution, we felt that participants tended to repeat the same themes, which satisfied our requirement for data saturation. Additionally, study results cannot be generalizable to other women at risk for hypertension, since the study took place in a specialized renal clinic located in an urban setting.

Conclusion

This study aimed to better understand the perceptions and needs of PHRPE and their health care providers with respect to telemonitoring. Through semistructured interviews with patients and providers, the benefits of, barriers to, and desired features of a telemonitoring program were identified. Patients and providers were hopeful about the benefits that such a program may provide in terms of self-care and clinical management of PHRPE. Perceived benefits included close monitoring of patients; tailored access to health care services, when needed; and a sense of empowerment for self-management. Perceived barriers to the telemonitoring of PHRPE included financial barriers as well as a potential increase in clinician workload. Desired features included the facilitation of self-management by patients, facilitation of decision making by providers, and provision of evidenced-based action prompts. Recommended features for a telemonitoring program for PHRPE were provided and were based on the perceived benefits, barriers, and desired features, as described by the patient and provider participants. Additionally, the use of theoretical frameworks in the design and implementation of a telemonitoring program for PHRPE, such as the empowerment-based approach for self-management and the person-based approach used to develop a logic model, were discussed as potential beneficial tools. The findings from this study validate the need for telemonitoring programs for PHRPE. The recommendations from this research may provide valuable insights into the development of future telemonitoring programs to improve self-care and the clinical management of PHRPE.

Authors' Contributions

MA contributed to the acquisition, analysis, and interpretation of data, as well as the design, draft, and revision of the work. JG contributed to the analysis and interpretation of data as well as the design and revision of the work. TV contributed to the interpretation of data as well as the draft and revision of the work. SM contributed to the design, analysis, and interpretation of data, as well as the revision of the work. MH contributed to the interpretation of data as well as the revision of the work. ES contributed to the design, analysis, and interpretation of data, as well as the conception, design, and revision of the work.

Conflicts of Interest

MH has received grants from Pfizer, Ionis, Chemocentryx, Calliditas and Roche; consultant fees from Alynylam; and royalties from UpToDate. The other authors have no conflicts of interest to disclose.

Multimedia Appendix 1
Interview guides.
[DOCX File , 32 KB - humanfactors_v9i1e32545_app1.docx]

References

1. American College of Obstetricians. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. Obstet Gynecol 2013 Nov;122(5):1122-1131. [doi: 10.1097/01.AOG.0000437382.03963.88] [Medline: 24150027]
2. Duley L. The global impact of pre-eclampsia and eclampsia. Semin Perinatol 2009 Jun;33(3):130-137. [doi: 10.1053/j.semperi.2009.02.010] [Medline: 19464502]
3. Duckitt K, Harrington D. Risk factors for pre-eclampsia at antenatal booking: systematic review of controlled studies. BMJ 2005 Mar 02;330(7491):565. [doi: 10.1136/bmj.38380.674340.e0]
4. Magee LA, Pels A, Helewa M, Rey E, von Dadelszen P, Magee L, et al. Diagnosis, evaluation, and management of the hypertensive disorders of pregnancy: executive summary. J Obstet Gynaecol Can 2014 May;36(5):416-438. [doi: 10.1016/s1701-2163(15)30588-0]
5. Wallis AB, Saftlas AF, Hsia J, Atrash HK. Secular trends in the rates of preeclampsia, eclampsia, and gestational hypertension, United States, 1987-2004. Am J Hypertens 2008 May 01;21(5):521-526. [doi: 10.1038/ajh.2008.20] [Medline: 18437143]
6. Bellamy L, Casas J, Hingorani AD, Williams DJ. Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. BMJ 2007 Nov 01;335(7627):974. [doi: 10.1136/bmj.39335.385301.be]
7. Khan KS, Wojdyla D, Say L, Gülmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: a systematic review. Lancet 2006 Apr;367(9516):1066-1074. [doi: 10.1016/s0140-6736(06)68397-9]
8. MacKay A. Pregnancy-related mortality from preeclampsia and eclampsia. Obstet Gynecol 2001;97(4):533-538. [doi: 10.1016/s0029-7844(00)01223-0]
9. Norwitz E, Bellussi F. Early Pregnancy Prediction of Preeclampsia. 2021. URL: https://www.uptodate.com/contents/early-pregnancy-prediction-of-preeclampsia [accessed 2021-01-25]
10. Brown MA. Is there a role for ambulatory blood pressure monitoring in pregnancy? Clin Exp Pharmacol Physiol 2014 Jan 30;41(1):16-21. [doi: 10.1111/1440-1681.12106] [Medline: 23651133]
11. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [FREE Full text] [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
12. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
13. Aquino M, Munce S, Griffith J, Pakosh M, Munnery M, Seto E. Exploring the use of telemonitoring for patients at high risk for hypertensive disorders of pregnancy in the antepartum and postpartum periods: scoping review. JMIR Mhealth Uhealth 2020 Apr 17;8(4):e15095 [FREE Full text] [doi: 10.2196/15095] [Medline: 32301744]
14. Hinton L, Tucker KL, Greenfield SM, Hodgkinson JA, Mackillop L, McCourt C, et al. Blood pressure self-monitoring in pregnancy (BuMP) feasibility study; a qualitative analysis of women's experiences of self-monitoring. BMC Pregnancy Childbirth 2017 Dec 19;17(1):427 [FREE Full text] [doi: 10.1186/s12884-017-1592-1] [Medline: 29258469]
15. Hinton L, Hodgkinson J, Tucker KL, Rozmovits L, Chappell L, Greenfield S, et al. Exploring the potential for introducing home monitoring of blood pressure during pregnancy into maternity care: current views and experiences of staff-a qualitative study. BMJ Open 2020 Dec 01;10(12):e037874 [FREE Full text] [doi: 10.1136/bmjopen-2020-037874] [Medline: 33262186]
16. Band R, Hinton L, Tucker KL, Chappell LC, Crawford C, Franssen M, et al. Intervention planning and modification of the BUMP intervention: a digital intervention for the early detection of raised blood pressure in pregnancy. Pilot Feasibility Stud 2019 Dec 20;5(1):153 [FREE Full text] [doi: 10.1186/s40814-019-0537-z] [Medline: 31890265]
17. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g]
18. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
19. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
20. Fetal Medicine Foundation. Risk Assessment: Risk for Preeclampsia. URL: https://fetalmedicine.org/research/assess/preeclampsia/first-trimester [accessed 2022-01-19]
21. 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]
22. Andrès E, Meyer L, Zulfiqar AA, Hajjam M, Talha S, Bahougne T, et al. Telemonitoring in diabetes: evolution of concepts and technologies, with a focus on results of the more recent studies. J Med Life 2019 Aug;12(3):203-214 [FREE Full text] [doi: 10.25122/jml-2019-0006] [Medline: 31666818]
23. Fletcher BR, Hinton L, Hartmann-Boyce J, Roberts NW, Bobrovitz N, McManus RJ. Self-monitoring blood pressure in hypertension, patient and provider perspectives: a systematic review and thematic synthesis. Patient Educ Couns 2016 Mar;99(2):210-219. [doi: 10.1016/j.pec.2015.08.026] [Medline: 26341941]

24. Alpay L, van der Boog P, Dumaij A. An empowerment-based approach to developing innovative e-health tools for self-management. Health Informatics J 2011 Dec;17(4):247-255 [FREE Full text] [doi: 10.1177/1460458211420089] [Medline: 22193825]

Abbreviations

BP: blood pressure
BUMP: Blood Pressure Self-Monitoring in Pregnancy
DHI: digital health intervention
PHRPE: patients at high risk for preeclampsia

Edited by A Kushniruk; submitted 02.08.21; peer-reviewed by K Tucker, J van den Heuvel; comments to author 02.10.21; revised version received 24.11.21; accepted 28.11.21; published 07.02.22.

Please cite as:
Aquino M, Griffith J, Vattaparambil T, Munce S, Hladunewich M, Seto E
Patients' and Providers' Perspectives on and Needs of Telemonitoring to Support Clinical Management and Self-care of People at High Risk for Preeclampsia: Qualitative Study
JMIR Hum Factors 2022;9(1):e32545
URL: https://humanfactors.jmir.org/2022/1/e32545
doi:10.2196/32545
PMID:35129445

Review

Opportunities and Challenges for Professionals in Psychiatry and Mental Health Care Using Digital Technologies During the COVID-19 Pandemic: Systematic Review

Hélène Kane[1], PhD; Jade Gourret Baumgart[1], MA; Wissam El-Hage[2,3], Prof Dr; Jocelyn Deloyer[4], MA; Christine Maes[4], BA; Marie-Clotilde Lebas[5], MA; Donatella Marazziti[6,7], MD; Johannes Thome[8], MD, PhD; Laurence Fond-Harmant[9,10], PhD, HDR; Frédéric Denis[1], HDR

[1]Laboratoire Éducation, Éthique, Santé, Université de Tours, Tours, France

[2]Centre d'Investigation Clinique, Institut National de la Santé et de la Recherche Médicale, Tours, France

[3]Centre Hospitalier Régional Universitaire Tours, Tours, France

[4]Centre Neuropsychiatrique St-Martin, Namur, Belgium

[5]Département des Sciences de la Santé Publique et de la Motricité, Haute Ecole de la Province de Namur, Namur, Belgium

[6]Department of Experimental and Clinical Medicine, Section of Psychiatry, University of Pisa, Pisa, Italy

[7]Dipartimento di Medicina Clinica e Sperimentale, Section of Psychiatry, University of Pisa, Unicamillus-Saint Camillus International University of Health Sciences, Rome and Brain Research Foundation, Lucca, Italy

[8]Department of Psychiatry, University of Rostock, Rostock, Germany

[9]Agence de Coopération Scientifique Europe-Afrique-Luxembourg, Luxembourg, Luxembourg

[10]Education et Pratiques en Santé, Université Sorbonne Paris Nord, Paris, France

Corresponding Author:
Hélène Kane, PhD
Laboratoire Éducation, Éthique, Santé
Université de Tours
Boulevard Tonnellé
Tours, 37032
France
Phone: 33 279060019
Email: helene.kane@gmail.com

Abstract

Background: The COVID-19 pandemic has required psychiatric and mental health professionals to change their practices to reduce the risk of transmission of SARS-CoV-2, in particular by favoring remote monitoring and assessment via digital technologies.

Objective: As part of a research project that was cofunded by the French National Research Agency (ARN) and the Centre-Val de Loire Region, the aim of this systematic literature review was to investigate how such uses of digital technologies have been developing.

Methods: This systematic review was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The search was carried out in the MEDLINE (ie, PubMed) and Cairn databases, as well as in a platform specializing in mental health, Ascodocpsy. The search yielded 558 results for the year 2020. After applying inclusion and exclusion criteria, first on titles and abstracts and then on full texts, 61 articles were included.

Results: The analysis of the literature revealed a heterogeneous integration of digital technologies, not only depending on countries, contexts, and local regulations, but also depending on the modalities of care. Notwithstanding these variations, the use of videoconferencing has developed significantly, affecting working conditions and therapeutic relationships. For many psychiatric and mental health professionals, the pandemic has been an opportunity to build up their experience of remote care and, thus, better identify the possibilities and limits of these digital technologies.

Conclusions: New uses of such technologies essentially consist of a transition from the classic consultation model toward teleconsultation and make less use of the specific potential of artificial intelligence. As professionals were not prepared for these uses, they were confronted with practical difficulties and ethical questions, such as the place of digital technology in care,

confidentiality and protection of personal data, and equity in access to care. The COVID-19 health crisis questions how the organization of health care integrates the possibilities offered by digital technology, in particular to promote the autonomy and empowerment of mental health service users.

(JMIR Hum Factors 2022;9(1):e30359) doi:10.2196/30359

KEYWORDS

COVID-19; e–mental health; professional practices; quality of care; telepsychiatry; videoconferencing

Introduction

The spread of digital technology in health systems is a major and irreversible phenomenon, a source of changes that are only just beginning. Initiated several decades ago in the field of psychiatry and mental health care, the development of digital technologies has been increasing for several years [1-3]. Teleconsultation has notably begun to be used in specific contexts, such as when access to health care is at stake for expatriates or people living in isolated areas, while remaining marginal [4,5]. Technical difficulties, concerns about confidentiality, and regulatory barriers are among the obstacles to the development of telepsychiatry [6]. The expansion of new technologies offers more and more possibilities, including sensors that are able to collect clinical data related to physical activity, stress, or sleep. Intelligent applications are able to detect changes in individual behavior and then analyze this data to assist in screening and monitoring mental illnesses. Not only do such technologies open up new possibilities, but they might also bring about decisive changes to enhance the overall efficiency of mental health services [1,3].

The COVID-19 pandemic has highlighted the potential of these technologies, which have led to digital uses on an unprecedented scale in psychiatry. In particular, the pandemic revealed the contributions of these technologies to ensuring continuity of care while annihilating the risk of viral transmission in the context of an outbreak. As they allow remote monitoring of some patients, these technologies have been used in a wide range of strategies to reduce the risk of transmission of SARS-CoV-2. They have also made it possible to carry out interventions responding to needs that are specifically related to the epidemic, whether it be support for frontline health professionals or care for patients with COVID-19. The use of teleconsultation, previously in mental health care and psychiatry in its early stages, has massively increased in response to the health crisis and among measures that have been implemented to contain it [7,8]. These experiences of telepsychiatry, which started in emergency situations and were facilitated by exceptional arrangements and, often, regulatory relaxation [9], raise many questions about the evolution of health care and involve ethical and regulatory issues [10].

Mental health and psychiatric care specifically provide a central place to the therapeutic relationship. In this context, our attention is focused on the impact of digital technologies as a “relational artifact“ (ie, the way they reconfigure care relationships).

The objectives of this study were (1) to describe the uses of digital technologies at the time of COVID-19 and their impact on professional practices in psychiatry and mental health and (2) to understand the place of digital technologies in the organizational adaptations linked to the COVID-19 epidemic, but also to identify how this specific context questions the modalities of care.

Methods

This systematic review was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [11].

Search Strategy

A systematic literature search was carried out in two databases, MEDLINE (ie, PubMed) and Cairn, and a specialized mental health platform, Ascodocpsy; all included articles met the inclusion criteria. Search terms were defined by articulating keywords, which were previously defined from dictionaries of synonyms and thesauri *alone* or *in combination with* the Boolean operators “AND” and ”OR” (Table 1). The search, which was carried out on titles and abstracts, only concerned the year 2020, with extraction of results taking place as of December 31, 2020. Our process did not generate any results in previous years, which is consistent with the date that the COVID-19 epidemic began. Only peer-reviewed articles accepted for publication and available in English or French were included. Editorials were excluded.

After a preliminary exploratory search, all authors agreed on the inclusion and exclusion criteria. To be included in the literature review, articles had to meet the following criteria: deal with the use of digital technologies as a response to the pandemic context and be related to the field of mental health care or psychiatry. On the other hand, the following were excluded: articles documenting the impacts of COVID-19 on mental health and psychiatry in general, adaptations of the health care offering carried out independently of the digital possibilities, and uses of digital tools in mental health and psychiatry independent of the COVID-19 outbreak.

In order to benefit from international experiences in an unprecedented context where many countries were simultaneously confronted with the same challenges, we chose not to exclude references based on geographical criteria.

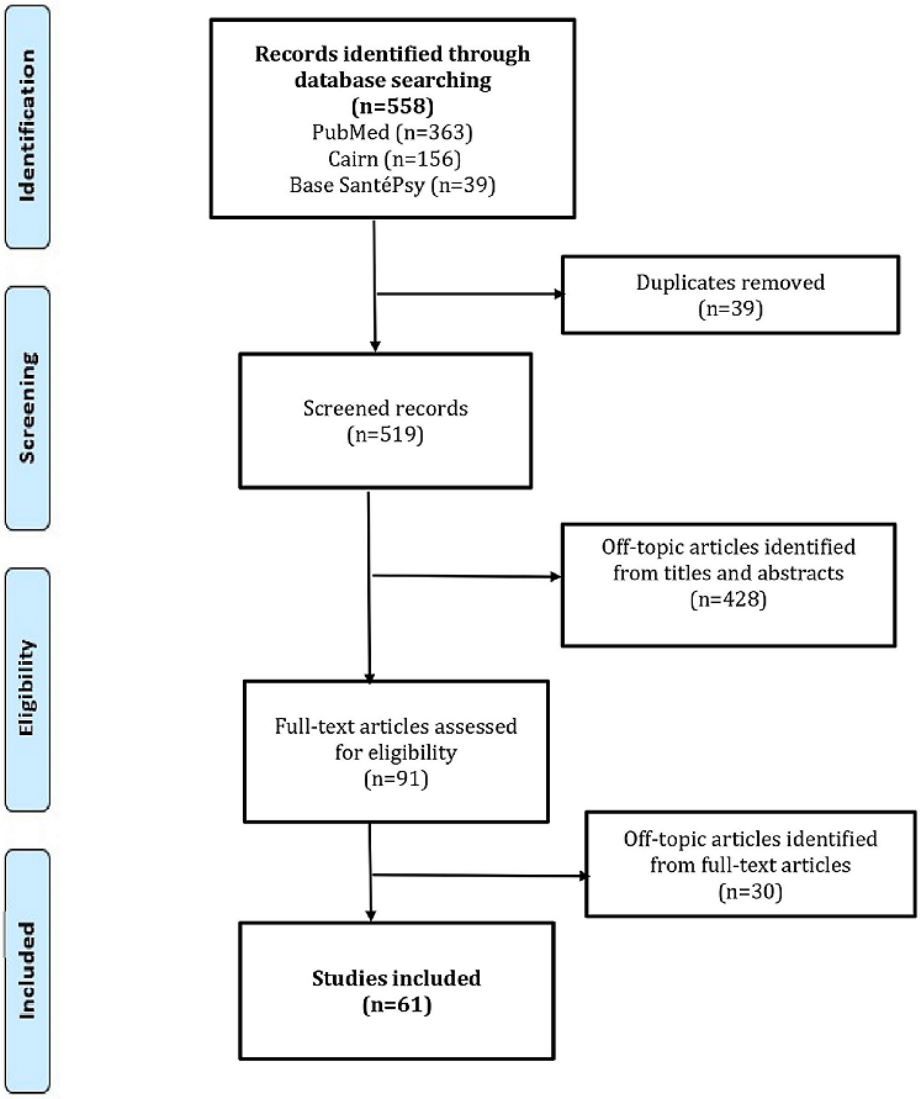
Table 1. Search terms used to find articles for this review.

Database	Thesaurus	Search terms
PubMed (MEDLINE)	Yes	(“coronavirus” OR “covid-19” OR “sars-cov-2”) AND (“mental health worker” OR “psychiatry” OR “mental health professional” OR “psychiatrist” OR “psychologist” OR “psychiatric nurse” OR “e-professional in psychiatry” OR “e-mental health”)
Cairn	No	(“covid-19” OU “sars-cov-2” OU “coronavirus”) ET (“psychiatrie” OU “santé mentale” OU “psychologue” OU “infirmier en psychiatrie” OU “pair-aidant” OU “médiateur de santé pair” OU “e-professionnel de la psychiatrie”)
Base SantéPsy (Ascodocpsy)	Yes	Base set contains “covid-19” ET (“psychiatrie” OU “santé mentale” OU “psychologie” OU “hôpital psychiatrique”)

Study Detection

The search yielded a total of 558 documents, 39 of which were duplicates that were excluded. The first two authors (HK and JGB) preselected references by applying the inclusion and exclusion criteria on abstracts and agreed to select 91 articles. A careful reading of the documents resulted in the exclusion, after consultation, of 30 more articles. Therefore, a total of 61 references were selected (Figure 1).

Figure 1. Flowchart of article selection for this review.



Results

Contributions From the Use of Digital Technologies in Mental Health Care and Psychiatry

This literature review shows that, in the context of the current crisis and as professional practices need to adapt, publications have been produced at a rapid rate. The use of digital technologies appeared to be a crucial issue, which was addressed in 61 articles in the year 2020 alone. Table 2 [5,7,12-70] describes the characteristics of these articles, the digital technologies they addressed, their main uses, and the fields of application mentioned.

The methodological quality of these contributions turned out to be quite poor, due to a lack of time and hindsight to carry out more rigorous work. About half of them (30/61, 49%) were feedback articles. They show the willingness of those involved in psychiatry and mental health care to share their experiences and innovations in the midst of the COVID-19 crisis. Such publications reflect an acceleration in the exchange of professional practices on an international scale. They used diverse methodologies, ranging from personal narratives to more collaborative and structured forms of feedback and analysis of experience. The presence of 8 (13%) reflection-based articles also shows the willingness of professionals to share their concerns. Another set of 14 (23%) articles were literature reviews, either narrative or systematic. In the end, of the 61 articles, only 9 (15%) were original research studies.

Selected articles included many countries, spread over five continents, which were simultaneously confronted with similar issues related to the use of digital technologies in response to the pandemic. The largest number of articles concerned Western and Northern Europe (n=24, 39%) and North America (n=23, 38%).

Table 2. Characteristics and themes of the selected articles.

Characteristics	Studies (N=61), n (%)	References
Type of article		
Experience feedback	30 (49)	[12-41]
Literature review	14 (23)	[5,7,42-53]
Study	9 (15)	[54-62]
Reflection	8 (13)	[63-70]
Location of study		
Europe	24 (39)	[5,13,14,18,19,26,27,29,31,33,36,38,40,42,44,50,54,57,58,60-62,64,65]
North America	23 (38)	[7,12,15,17,20-24,30,32,34,35,37,39,41,45,46,48,49,52,63,66]
Asia	8 (13)	[16,47,51,53,55,67,68,70]
Australia	4 (7)	[25,43,59,69]
Africa	2 (3)	[28,56]
Digital tool		
Videoconferencing	45 (74)	[22,24,26-35,37,39-41,47,50,52,58,59,61,66,68-70]
Telephone	27 (44)	[24-27,32,33,35-37,39,51,59,60,62,68,69]
App	10 (16)	[29,47-49,51,53-55,64,67]
Connected device	5 (8)	[47-49,65,67]
Website	4 (7)	[30,53,56,68]
Artificial intelligence	4 (7)	[48,49,64,67]
Use of digital tools		
Patients' follow-up and care	44 (72)	[5,7,12-21,23,25-29,31-38,41-43,46,48-50,52,57,58,60,61,63,65-68,70]
Public support for COVID-19	9 (15)	[35,44,49,51,53,54,56,62,64]
Group therapy	8 (13)	[17,19,21,30,34,39,50,67]
Assessment and diagnosis	7 (11)	[22,25,45,49,60,62,65]
Support for health professionals	5 (8)	[30,39,44,55,62]
Care of patients with COVID-19	5 (8)	[14,16,17,40,53]
Staff meeting	5 (8)	[7,16,27,34,41]
Domain		
Psychiatry in general	29 (48)	[5,7,14,15,17,20,21,24,25,28,34,40-44,46,50,52,55,57-60,63,66-69]
Psychology	14 (23)	[16,24,30,33,35,36,39,40,47,49,53,54,62,64]
Mental health promotion	8 (13)	[44,47,49,51,53,56,62,64]
Child psychiatry	5 (8)	[13,18,29,32,37]
Geriatric psychiatry	5 (8)	[31,38,56,61,65]
Community health	4 (7)	[12,23,27,68]
Forensic psychiatry	3 (5)	[22,26,45]
Addiction	3 (5)	[19,48,70]

Adopting New Digital Technologies

Most of the articles selected for this literature review (45/61, 74%) mentioned the use of videoconferencing. This technology has been used, in particular, in interventions with mental health and psychiatric professionals who provide care for patients with COVID-19 [30,39,55] or, more broadly, for the population affected by the COVID-19 pandemic [40,51,62]. Videoconferencing consultations and group discussions took place alongside on-site interventions and telephone helplines.

However, teleconsultation has been massively developed to ensure continuity of care. Videoconferencing has made it possible to maintain not only remote monitoring of patients, but also therapy groups [17,48] and meetings between professionals. This particular use of this technology at an unprecedented scale has been developed at the intersection of

two phenomena: the constraint of physical distancing measures aimed at containing the epidemic and the achievement of a high level of digital performance that allows for seamless use of videoconferencing. This rapid expansion of teleconsultation is perhaps the most important impact of the pandemic on the organization of care in psychiatry [7]. This sudden evolution has often gone along with the use of videoconferencing platforms, such as Skype, Zoom, or Microsoft Teams, despite their use raising security and confidentiality issues [17,20,21,32,41,43]. Specialized digital health platforms, such as MyChart by Epic, have been less frequently employed. In some cases, the use of videoconferencing has been combined with that of more familiar telecommunication tools, such as the telephone [17,36,38,57,59], emails, or SMS [34]. In our literature review, the telephone was the second most frequently mentioned digital technology, with 27 out of 61 (44%) references cited.

In most psychiatric services, however, this switch to remote communication is not yet complete, forcing practitioners to determine which activities require face-to-face meetings and which ones can be done via videoconferencing [14,16,23,25]. In particular, in-person examination has been maintained for patients who are deemed vulnerable and at risk [25].

As the number of remote consultations increased, prescription procedures have also been impacted. In order to limit the number of in-person appointments, practitioners have either used tele-prescription or opted for prescriptions covering a longer period. For treatments that require follow-up of specific clinical parameters, such as Clozapine, which involves monitoring blood counts, protocols have been made more flexible, sometimes allowing for a remote assessment of the clinical condition of patients [23,29,71].

In addition, 10 articles (16%) discussed connected apps and devices. In particular, their authors highlighted the relevance of connected apps and devices to assist remote monitoring during the pandemic [48,54,64]. Although the analyzed literature mentioned the potential of these new technologies, none of the identified publications documented the possible increase in their use in the context of a health crisis. Only 2 (3%) references were about apps that enabled the connection between patients and health professionals [51,55].

Heterogeneous Integration of Digital Technologies

Although the COVID-19 pandemic has stimulated the development of telepsychiatry on all continents, our review of the literature allowed us to glimpse variations between countries. In the United States, due to the removal of regulatory barriers, the shift to telepsychiatry has been massive and even total in certain units, as illustrated by numerous publications [15,24,32,34,37,41]. In many European countries, telepsychiatry has gone from a niche practice to an essential modality for providing mental health services; this has been so in Germany [27], Spain [62], France [13,18,42], Ireland [19,29,50], the Netherlands [57], and Switzerland [65]. Telemedicine was adopted, even in countries such as North Macedonia, where public policies had previously been rather opposed to it [58]. A similar situation was reported in Australia [25,43,69]. In some cases, barriers persist, as in Canada, where the lack of health insurance coverage for teleconsultation with mental health professionals prevented its expansion during the pandemic [35]. Other countries, such as India, have been quick to innovate in favor of integrating digital technologies [47,68], and they rely on the development of telepsychiatry in order to increase health care delivery, despite limited resources. Developing telemedicine has been more difficult in some countries of the Global South due to a lower spread of information and communication technology [28,51]. Although advancing heterogeneously, the COVID-19 pandemic has been stimulating the integration of digital technologies into health care, confronting many countries simultaneously with comparable problems.

Within countries, these trends raise the issue of unequal access to digital technologies. Consequently, the development of telepsychiatry may disadvantage people living in poverty [23,63] and older adults who do not have access to these technologies [31,38,56].

Moreover, many authors have reassessed the appropriateness of telepsychiatry depending on the patients and their disorders, which had already been documented in the literature [3]. Telepsychiatry seems inappropriate for use with children whose attention is difficult to capture from a distance [29,69], as well as with older patients, who are less familiar with digital technologies and frequently suffer from hearing and visual impairments [38,60,61]. On the other hand, telepsychiatry is likely to facilitate access to health care for youth who are accustomed to new technologies [32].

Specific problems with the use of these technologies have arisen in certain fields, such as forensic psychiatry [45,72], the treatment of drug addiction [70], or electrostimulation techniques [7]. Although facilitating some procedures [22], remote forensic assessments are at risk of being disqualified for “procedural defects” [45]. On the other hand, the possibility of appearing in court by videoconference has prevented some forms of stress for people with mental disorders [26]. With regard to the treatment of drug addiction, specific difficulties may be related to legislative measures taken to prevent the diversion and misuse of certain drugs, by prohibiting tele-prescription, among other restrictions [70].

These numerous contributions found in the scientific literature, which were based on new experiments in the context of the COVID-19 pandemic, have added to the established knowledge about the relevance of telepsychiatry in different situations.

Experiencing New Conditions of Professional Practice

The use of telepsychiatry, which makes it possible to reduce the risks of infection, has generated new conditions of practice for many professionals, defining both new possibilities and constraints. For independent practitioners, teleconsultation is no longer necessarily a freely chosen practice [4], but a means of maintaining their practice despite the restrictions imposed by public health measures. In hospital services, the decision to switch some services to a remote mode has been taken by those in charge in a more or less constrained way, or more or less consensually. Many professionals have had to adapt their

practices, even though they were initially hostile to the use of digital technologies [5,42].

It should be noted that this new digital work experience has sometimes been associated with teleworking from home [27,33], in a general context of lack of preparation. Most of the professionals concerned were not trained in telepsychiatry follow-up, and some were not very comfortable with new technologies [33]. At the same time, they had to learn how to use digital technologies to support their patients, manage their own stress, and sometimes set up home-based work processes [15,52]. The accumulation of all these tensions can lead to emotional exhaustion [52].

Professionals teleworking from home have been confronted with unprecedented situations of temporal and spatial juxtaposition of both their professional and personal lives. This juxtaposition requires "psychological work to differentiate between private and professional lives that is more costly than usual" [33]. However, other authors mention that telework can also facilitate work-life balance in psychiatry [41,60], and practitioners reported that caring for a patient while teleworking nevertheless had a positive impact on their well-being in the midst of the crisis [41].

Moreover, the use of digital technologies, especially as their use is improvised and unframed, is likely to lead to an increased workload. Professionals may be exposed to an accumulation of requests through multiple technologies: videoconferencing, telephone, email, and SMS [15]. For them, the extensive use of videoconferencing can be a source of fatigue [24,52], described as "Zoom fatigue" [34], and a source of stress [34,52]. This fatigue is related to efforts to communicate and establish a relationship via videoconferencing [15], to the difficulty of sticking to a schedule with time slots that are explicitly dedicated to each patient, and to all the operations required to disconnect and reconnect to each device [34]. And yet, after overcoming technical and organizational obstacles, professionals can take advantage of those digital technologies, which can also bring more flexibility and help them save time [57], as documented in research prior to the pandemic [73].

Having experiences of care relationships reshaped by digital technologies in the context of this pandemic, psychiatric and mental health professionals have been using videoconferencing and the telephone to follow up on many patients. For health care providers, the COVID-19 crisis has been an opportunity to build up their experience of remote health care monitoring and, thus, better understand the possibilities and limitations of such digital technologies. This unprecedented context forced them to reinvent "relational mental health" [42]. Among other consequences, telepsychiatry favors more fragmented care modalities, with shorter and more frequent encounters [15,57].

Furthermore, although the effectiveness of telepsychiatry had already been documented [3,50,74], many professionals were skeptical about the possibility of establishing, maintaining, and strengthening a therapeutic relationship [66]. They feared that they would lose not only human contact, but also control over their image [3]. Some feared that the screen would become a barrier to the therapeutic process [29]. After experiencing telepsychiatry in the context of the health crisis, assessments remain contrasted [20]. Some authors mentioned that teleconsultation tends to hinder verbal and nonverbal communication [20,30,34,48,53]. In addition to the difficulty in grasping nonverbal body language [20,34,53], it is no longer possible to smell odors [46,53], see how patients are dressed, and perceive certain attitudes [46]. Communication must do without physical contact, such as greeting each other with a handshake [29], and empathy can no longer be manifested by comforting gestures [61]. Videoconferencing introduces a new mode of presence to the other, sometimes inducing a feeling of dissonance due to audiovisual presence and bodily absence [20]. The feeling of intimacy and confidence is not the same as in a closed office, and consultation no longer benefits from a separate space and time but is embedded in everyday life [30].

The professionals were also led to discover the advantages of digital technologies. The use of videoconferencing can be an opportunity to better contextualize some information, since part of the patient's environment is made visible [16,53]. Teleconsultation can also make it possible to remove certain inhibitions and to access the unconscious more easily [9]. A few authors identified advantages of telephone consultations over videoconferences, especially for short talks [55]. In some cases, conducting telephone consultations allowed patients greater freedom of expression and allowed professionals to listen more carefully [14].

Although telepsychiatry allows for a large number of follow-up consultations, several authors mentioned that the greatest difficulty was in establishing a therapeutic relationship without a prior face-to-face encounter [46]. As a result, in the context of COVID-19, psychiatric and mental health professionals tended to postpone work on trauma, focusing instead on maintaining the patient's well-being and encouraging activities to achieve this goal [34]. With regard to diagnosis, almost all professionals interviewed in a study that was conducted in Ireland reported that they were less comfortable making a diagnosis based on a telephone consultation [56]. Similarly, conducting neuropsychological assessments from a distance presents specific difficulties [14,61].

Although already documented, the possibilities and limits of telepsychiatry were highlighted by the COVID-19 epidemic, illustrating how experience can help to gradually shape new therapeutic practices integrating digital technologies.

Discussion

Principal Findings

The profusion of articles identified in the framework of this literature review shows how much the COVID-19 crisis has raised issues about care practices in psychiatry and how they integrate the available digital technologies. Such integration proves to be heterogeneous, depending on local contexts and regulations, but also regarding the fields and modalities of care. The use of videoconferencing has had an impact not only on the working conditions of mental health and psychiatric professionals, but also on the care relationships they maintain with their patients. This sudden shift to remote care has prompted professionals to publish papers about their experiences

with telepsychiatry, sometimes in a naive way, without building on pre-existing research.

Lack of Preparation Confronting Professionals With Ethical Questions

The experience of videoconferencing, in a context where mental health and psychiatric professionals had not been prepared for it, calls into question the quality of care [60]. Many professionals have used teleconsultation without training or knowledge of existing protocols and recommendations [75]. This lack of preparation pushed them to improvise and confronted them with ethical dilemmas. This unprecedented situation raises questions about privacy and the protection of personal data [47] as well as the risk of increasing inequalities in access to health care [48,50,63]. Indeed, the rapid expansion of telemedicine hinders access to health care for patients who do not possess, nor are proficient in, the necessary technology [9].

To guarantee the best conditions for confidentiality, special attention should be paid to the choice of digital technologies to be used. In the context of the current crisis, this choice has been little considered and is essentially based on pragmatic considerations. The use of new technologies requires special precautions, such as using headphones, consulting in a closed room, and disconnecting when absent [20]. Therefore, their use requires awareness and support from professionals in order to ensure the protection of personal data [47]. In addition, remote intervention involves knowing where the patient is and what local resources are available to respond to emergency situations [75].

The partial or total shift to remote consultations also raises the issue of equity in the provision of health care. Many professionals have been forced to identify vulnerable patients who require face-to-face encounters and those who can be monitored remotely [14,15,23,25]. This process of triage and separation between patients has confronted them with ethical dilemmas, particularly in the case of patients who are at risk of both a relapse of their mental illness and developing severe forms of COVID-19. Guidelines were sometimes developed to help identify patients for whom a face-to-face appointment was absolutely necessary [12]. Some authors warned against excluding vulnerable people or people living in poverty who do not have access to the internet or are limited by low digital literacy [63,68]. The situation of older adult patients, who are often unfamiliar with new technologies, also requires special attention. Not supporting them in the use of these tools constitutes a form of ageism [31]. Not only are older people not always averse to new technologies, but the issue of distance is all the more important as they are vulnerable to COVID-19 [65]. The experience of these new uses of digital technology opens the way to forms of differentiated care that make it possible to adapt care delivery to patients' preferences, in order to improve the overall quality of care.

Uses That Do Not Exploit the Full Potential of Digital Technology

In showing many professionals the potential of digital technologies, the COVID-19 crisis also revealed the extent to which their nonuse can be an ethical challenge. Digital technologies can reduce regional inequalities in access to health care [52] and can prevent significant costs and travel time [18,68]. Moreover, many patients have expressed their satisfaction with their experience of telepsychiatry [19,32,46,58], which confirms previous data from the literature [3]. Although some patients feel less supported in teleconsultations, others appreciate the freedom to access health care from an online platform [19]. Telepsychiatry can also foster patient autonomy [19,50,76] and the development of a form of empowerment that health organizations have been advocating for [77]. For example, it allows patients to record consultations in order to fully assimilate the information, thus increasing their power to act on the care relationship [50]. The experience of videoconferencing has allowed some professionals to overcome part of their initial reluctance [6] and to realize how valuable the virtual space of teleconsultation can be for building certain forms of intimacy in the therapeutic relationship [76,78]. However, in emergency and unpreparedness contexts, the experience of telepsychiatry did not take place in optimal conditions for their appropriation. In particular, some professionals experimented with videoconferencing when they were teleworking and not in their offices. The COVID-19 crisis has, nevertheless, made visible the extent to which digital technology can be a driving force for change in psychiatry [26]. It paves the way for the development of a hybrid care system integrating the strengths of teleconsultations as a complement to face-to-face encounters [3,8,76].

However, while the use of teleconsultation has been significant, professional uses of apps and connected devices do not seem to have been as stimulated by the health crisis. According to some authors, the pandemic, nevertheless, made it urgent to use such tools in order to intervene on a large scale to relieve the mental health burden induced by the crisis [54,67,79]. By providing clinical information, dedicated apps can also help to develop more personalized care plans [67]. The possibilities offered by digital phenotyping open up new perspectives for remote monitoring and assessment, making it possible, in particular, to detect the occurrence of disorders or relapses [48,49]. Even in a crisis context, the use of these tools has been hampered by the lack of both evaluations proving their effectiveness [54,64,79] and appropriate regulations [64]. Although the use of digital technologies has been stimulated by the crisis, in health care systems, it has essentially consisted of a transfer from in-person consultation to teleconsultation. In the end, health care has not taken full advantage of the specific intervention potential of digital technologies and artificial intelligence.

Challenges to Better Integration of Digital Technologies in the Organization of Health Care

The COVID-19 health crisis questions the organization of care and the way it integrates new possibilities offered by digital technology. The scientific publications that we have identified mainly addressed the issues related to teleconsultation, sometimes ignoring previous research. The impact of the use of digital technology on relationships between professionals is poorly documented. However, the use of digital technologies is reshaping the conditions of teamwork and allows for new modalities of interprofessional collaboration [80,81]. The stakes

are high, since digital technologies open up the possibility of numerous contacts between professionals working in different places. In particular, they make possible new collaborations between mental health and somatic professionals, and they allow a reorganization of consultation-liaison psychiatry. Constraints also need to be examined, since videoconference work meetings reduce interactional diversity, especially informal interactions. Such development is likely to affect well-being at work and social support between professionals in the context of crisis. However, other research has shown that telepsychiatry can allow professionals to optimize their working time and reduce their risk of burnout [73]. Remote team management, however, requires specific approaches [80].

Digital technologies also question the place of users and their relatives in the organization of care. As a result of the increase in outpatient and remote follow-up during lockdown periods, many patients have become more autonomous in managing their mental health [50]. Telepsychiatry can support the development of integrated patient-centered care, allowing for a more precise match between health care providers' skills and patients' needs [81]. Families have also been placed in the front line, pushed to take on new responsibilities [18,68]. Because community health services limited their travel, families relied more heavily on community resources [23]. New possibilities offered by digital technology thus invite new research to be conducted into community-based approaches to mental health.

Conclusions

The COVID-19 pandemic has led to new uses of telepsychiatry, with the aim of ensuring continuity of care while limiting the risk of transmission of SARS-CoV-2. Such expansion was essentially characterized by the integration of videoconferencing as a new framework for consultation. Many mental health and psychiatric professionals started experiencing remote health care monitoring and assessment in a hurry and with no preparation. They have become familiar with the constraints, possibilities, and assets of care relationships in this type of context. These new conditions of professional practice have confronted them with ethical questions, such as equity in access to care. Existing research resources and data could be mobilized to enable these professionals to better leverage the benefits of digital technologies to complement face-to-face meetings. Further interdisciplinary work will be needed to better understand variations in digital technology uses across countries.

The use of digital technologies during the COVID-19 epidemic have shed light on the organization of mental health and psychiatric care, and about the place of users within this context of care. In a context where hospitals and health centers are no longer the only spaces where care and support are delivered, access to care and "decoding" the eHealth world constitute a pillar of tomorrow's public health [71].

Acknowledgments

We thank the French National Research Agency (ARN) and the Centre-Val de Loire Region for funding our research.

Conflicts of Interest

WEH received speaker fees from Air Liquide, Eisai, Janssen, Lundbeck, Otsuka, UCB Pharma, and Chugai. WEH has received research grants from the Fondation de France and from the French National Hospital Program for Clinical Research (PHRC) that are unrelated to the submitted work.

References

1. Briffault X. Psychiatrie 3.0: Etre Soi et ses Connexions. Paris, France: Doin; 2019:1-223.
2. Shore J. The evolution and history of telepsychiatry and its impact on psychiatric care: Current implications for psychiatrists and psychiatric organizations. Int Rev Psychiatry 2015;27(6):469-475. [doi: 10.3109/09540261.2015.1072086] [Medline: 26397182]
3. Chakrabarti S. Usefulness of telepsychiatry: A critical evaluation of videoconferencing-based approaches. World J Psychiatry 2015 Sep 22;5(3):286-304 [FREE Full text] [doi: 10.5498/wjp.v5.i3.286] [Medline: 26425443]
4. Mathieu-Fritz A. Les téléconsultations en santé mentale. Ou comment établir la relation psychothérapeutique à distance. Réseaux 2018;207(1):123-126. [doi: 10.3917/res.207.0123]
5. Advenier F, Reca M. Téléconsultations pendant le confinement en cabinet de ville. Inf Psychiatr 2020;96(8):601-607 [FREE Full text]
6. Chan S, Parish M, Yellowlees P. Telepsychiatry today. Curr Psychiatry Rep 2015 Nov;17(11):89. [doi: 10.1007/s11920-015-0630-9] [Medline: 26384338]
7. Cabrera MA, Karamsetty L, Simpson SA. Coronavirus and its implications for psychiatry: A rapid review of the early literature. Psychosomatics 2020;61(6):607-615 [FREE Full text] [doi: 10.1016/j.psym.2020.05.018] [Medline: 32943211]
8. Mishkind M, Shore J, Schneck C. Telemental health response to the COVID-19 pandemic: Virtualization of outpatient care now as a pathway to the future. Telemed J E Health 2021 Jul;27(7):709-711. [doi: 10.1089/tmj.2020.0303] [Medline: 33297849]
9. Kinoshita S, Cortright K, Crawford A, Mizuno Y, Yoshida K, Hilty D, et al. Changes in telepsychiatry regulations during the COVID-19 pandemic: 17 countries and regions' approaches to an evolving healthcare landscape. Psychol Med 2020 Nov 27:1-8. [doi: 10.1017/s0033291720004584]

10. CCNE. Numérique et Santé: Quels Enjeux Éthiques Pour Quelles Régulations?. Paris, France: CCNE; 2018 Nov 19. URL: https://www.ccne-ethique.fr/sites/default/files/publications/rapport_numerique_et_sante_19112018.pdf [accessed 2022-01-26]
11. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
12. Alavi Z, Haque R, Felzer-Kim IT, Lewicki T, Haque A, Mormann M. Implementing COVID-19 mitigation in the community mental health setting: March 2020 and lessons learned. Community Ment Health J 2021 Jan;57(1):57-63 [FREE Full text] [doi: 10.1007/s10597-020-00677-6] [Medline: 32681409]
13. Alibert JF. Les psychothérapies d'enfant en visio pendant le confinement: Pour le meilleur et pour le pire… Retour d'expérience. Inf Psychiatr 2020;96(8-9):617-622. [doi: 10.1684/ipe.2020.2161]
14. Barry H, Doherty AM, Clancy M, Moore S, MacHale S. Lockdown logistics in consultation-liaison psychiatry. Ir J Psychol Med 2020 Sep 11:1-9 [FREE Full text] [doi: 10.1017/ipm.2020.104] [Medline: 32912345]
15. Chen JA, Chung W, Young SK, Tuttle MC, Collins MB, Darghouth SL, et al. COVID-19 and telepsychiatry: Early outpatient experiences and implications for the future. Gen Hosp Psychiatry 2020;66:89-95 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.07.002] [Medline: 32750604]
16. Cheng W, Zhang F, Hua Y, Yang Z, Liu J. Development of a psychological first-aid model in inpatients with COVID-19 in Wuhan, China. Gen Psychiatr 2020;33(3):e100292 [FREE Full text] [doi: 10.1136/gpsych-2020-100292] [Medline: 32596642]
17. Childs AW, Unger A, Li L. Rapid design and deployment of intensive outpatient, group-based psychiatric care using telehealth during coronavirus disease 2019 (COVID-19). J Am Med Inform Assoc 2020 Jul 01;27(9):1420-1424 [FREE Full text] [doi: 10.1093/jamia/ocaa138] [Medline: 32687151]
18. Chokron S. Apport de la téléconsultation au bilan neuropsychologique de l'enfant. Rev Neuropsychol 2020;12(2):152-154. [doi: 10.1684/nrp.2020.0556]
19. Columb D, Hussain R, O'Gara C. Addiction psychiatry and COVID-19: Impact on patients and service provision. Ir J Psychol Med 2020 Sep;37(3):164-168 [FREE Full text] [doi: 10.1017/ipm.2020.47] [Medline: 32434597]
20. Cowan A, Johnson R, Close H. Telepsychiatry in psychotherapy practice. Innov Clin Neurosci 2020 Apr 01;17(4-6):23-26 [FREE Full text] [Medline: 32802589]
21. Datta N, Derenne J, Sanders M, Lock JD. Telehealth transition in a comprehensive care unit for eating disorders: Challenges and long-term benefits. Int J Eat Disord 2020 Nov;53(11):1774-1779. [doi: 10.1002/eat.23348] [Medline: 32715512]
22. Green AS, Ruchman SG, Katz CL, Singer EK. Piloting forensic tele-mental health evaluations of asylum seekers. Psychiatry Res 2020 Sep;291:113256 [FREE Full text] [doi: 10.1016/j.psychres.2020.113256] [Medline: 32619825]
23. Guan I, Kirwan N, Beder M, Levy M, Law S. Adaptations and innovations to minimize service disruption for patients with severe mental illness during COVID-19: Perspectives and reflections from an assertive community psychiatry program. Community Ment Health J 2021 Jan;57(1):10-17 [FREE Full text] [doi: 10.1007/s10597-020-00710-8] [Medline: 32930903]
24. Ingram DH, Best K. The psychodynamic psychiatrist and psychiatric care in the era of COVID-19. Psychodyn Psychiatry 2020;48(3):234-258. [doi: 10.1521/pdps.2020.48.3.234] [Medline: 32996853]
25. Kavoor A, Chakravarthy K, John T. Remote consultations in the era of COVID-19 pandemic: Preliminary experience in a regional Australian public acute mental health care setting. Asian J Psychiatr 2020 Jun;51:102074 [FREE Full text] [doi: 10.1016/j.ajp.2020.102074] [Medline: 32294583]
26. Kennedy H, Mohan D, Davoren M. Forensic psychiatry and Covid-19: Accelerating transformation in forensic psychiatry. Ir J Psychol Med 2021 Jun;38(2):145-153 [FREE Full text] [doi: 10.1017/ipm.2020.58] [Medline: 32434610]
27. Lyne J, Roche E, Kamali M, Feeney L. COVID-19 from the perspective of urban and rural general adult mental health services. Ir J Psychol Med 2020 Sep;37(3):181-186 [FREE Full text] [doi: 10.1017/ipm.2020.62] [Medline: 32434599]
28. Marehin M, Mboumba Hinnouo A, Obiang P. Organisation of psychiatric care in Gabon during the COVID-19 epidemic [Article in French]. Ann Med Psychol (Paris) 2021 Feb;179(2):137-140 [FREE Full text] [doi: 10.1016/j.amp.2020.08.015] [Medline: 32863389]
29. McGrath J. ADHD and Covid-19: Current roadblocks and future opportunities. Ir J Psychol Med 2020 Sep;37(3):204-211 [FREE Full text] [doi: 10.1017/ipm.2020.53] [Medline: 32434606]
30. Mellins C, Mayer L, Glasofer D, Devlin M, Albano A, Nash S, et al. Supporting the well-being of health care providers during the COVID-19 pandemic: The CopeColumbia response. Gen Hosp Psychiatry 2020;67:62-69 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.08.013] [Medline: 33059217]
31. Naarding P, Oude Voshaar RC, Marijnissen R. COVID-19: Clinical challenges in Dutch geriatric psychiatry. Am J Geriatr Psychiatry 2020 Aug;28(8):839-843 [FREE Full text] [doi: 10.1016/j.jagp.2020.05.019] [Medline: 32565007]
32. Ramtekkar U, Bridge JA, Thomas G, Butter E, Reese J, Logan E, et al. Pediatric telebehavioral health: A transformational shift in care delivery in the era of COVID-19. JMIR Ment Health 2020 Sep 18;7(9):e20157 [FREE Full text] [doi: 10.2196/20157] [Medline: 32525485]
33. Robin-Poupard F. Distanciation sociale, masques et visioconférence: Être thérapeute au temps de la Covid-19. J Psychol 2020;382(10):70-74. [doi: 10.3917/jdp.382.0070]

34. Sasangohar F, Bradshaw MR, Carlson MM, Flack JN, Fowler JC, Freeland D, et al. Adapting an outpatient psychiatric clinic to telehealth during the COVID-19 pandemic: A practice perspective. J Med Internet Res 2020 Oct 01;22(10):e22523 [FREE Full text] [doi: 10.2196/22523] [Medline: 32936768]
35. Scharf D, Oinonen K. Ontario's response to COVID-19 shows that mental health providers must be integrated into provincial public health insurance systems. Can J Public Health 2020 Aug;111(4):473-476 [FREE Full text] [doi: 10.17269/s41997-020-00397-0] [Medline: 32767269]
36. Scheider F. Le soutien téléphonique des patients isolés souffrant de troubles psychiques: Trame pour l'approche clinique en réhabilitation. Inf Psychiatr 2020:239-242. [doi: 10.1684/ipe.2020.2093]
37. Sharma A, Sasser T, Schoenfelder Gonzalez E, Vander Stoep A, Myers K. Implementation of home-based telemental health in a large child psychiatry department during the COVID-19 crisis. J Child Adolesc Psychopharmacol 2020 Sep;30(7):404-413. [doi: 10.1089/cap.2020.0062] [Medline: 32639849]
38. Usman M, Fahy S. Coping with the COVID-19 crisis: An overview of service adaptation and challenges encountered by a rural Psychiatry of Later Life (POLL) team. Ir J Psychol Med 2021 Dec;38(4):288-292 [FREE Full text] [doi: 10.1017/ipm.2020.86] [Medline: 32611473]
39. Viswanathan R, Myers MF, Fanous AH. Support groups and individual mental health care via video conferencing for frontline clinicians during the COVID-19 pandemic. Psychosomatics 2020;61(5):538-543 [FREE Full text] [doi: 10.1016/j.psym.2020.06.014] [Medline: 32660876]
40. Wulfman R, Jourdain P, Ourahou O. Pendant le trauma: Une approche de la pathologie psychiatrique des patients atteints par le Covid-19 à travers la plateforme Covidom. Inf Psychiatr 2020;96(8):627-631. [doi: 10.1684/ipe.2020.2163]
41. Yellowlees P, Nakagawa K, Pakyurek M, Hanson A, Elder J, Kales HC. Rapid conversion of an outpatient psychiatric clinic to a 100% virtual telepsychiatry clinic in response to COVID-19. Psychiatr Serv 2020 Jul 01;71(7):749-752. [doi: 10.1176/appi.ps.202000230] [Medline: 32460683]
42. Bocher R, Jansen C, Gayet P, Gorwood P, Laprévote V. Responsiveness and sustainability of psychiatric care in France during COVID-19 epidemic [Article in French]. Encephale 2020 Jun;46(3S):S81-S84 [FREE Full text] [doi: 10.1016/j.encep.2020.05.004] [Medline: 32471705]
43. Chherawala N, Gill S. Up-to-date review of psychotherapy via videoconference: Implications and recommendations for the RANZCP Psychotherapy Written Case during the COVID-19 pandemic. Australas Psychiatry 2020 Oct;28(5):517-520. [doi: 10.1177/1039856220939495] [Medline: 32689815]
44. Ćosić K, Popović S, Šarlija M, Kesedžić I. Impact of human disasters and COVID-19 pandemic on mental health: Potential of digital psychiatry. Psychiatr Danub 2020;32(1):25-31 [FREE Full text] [doi: 10.24869/psyd.2020.25] [Medline: 32303026]
45. Drogin EY. Forensic mental telehealth assessment (FMTA) in the context of COVID-19. Int J Law Psychiatry 2020;71:101595 [FREE Full text] [doi: 10.1016/j.ijlp.2020.101595] [Medline: 32768117]
46. Gautam M, Thakrar A, Akinyemi E, Mahr G. Current and future challenges in the delivery of mental healthcare during COVID-19. SN Compr Clin Med 2020 Jun 11:1-6 [FREE Full text] [doi: 10.1007/s42399-020-00348-3] [Medline: 32838140]
47. Grover S, Sarkar S, Gupta R. Data handling for e-mental health professionals. Indian J Psychol Med 2020 Oct;42(5 Suppl):85S-91S [FREE Full text] [doi: 10.1177/0253717620956732] [Medline: 33354071]
48. Hsu M, Ahern DK, Suzuki J. Digital phenotyping to enhance substance use treatment during the COVID-19 pandemic. JMIR Ment Health 2020 Oct 26;7(10):e21814 [FREE Full text] [doi: 10.2196/21814] [Medline: 33031044]
49. Melcher J, Hays R, Torous J. Digital phenotyping for mental health of college students: A clinical review. Evid Based Ment Health 2020 Nov;23(4):161-166. [doi: 10.1136/ebmental-2020-300180] [Medline: 32998937]
50. O'Brien M, McNicholas F. The use of telepsychiatry during COVID-19 and beyond. Ir J Psychol Med 2020 Dec;37(4):250-255 [FREE Full text] [doi: 10.1017/ipm.2020.54] [Medline: 32434596]
51. Roy A, Singh AK, Mishra S, Chinnadurai A, Mitra A, Bakshi O. Mental health implications of COVID-19 pandemic and its response in India. Int J Soc Psychiatry 2021 Aug;67(5):587-600 [FREE Full text] [doi: 10.1177/0020764020950769] [Medline: 32873106]
52. Schroeder R. Adaptation or revolution: Telemental health and advanced practice psychiatric nursing during COVID-19. J Am Psychiatr Nurses Assoc 2020 Nov 09:1-8 [FREE Full text] [doi: 10.1177/1078390320970638] [Medline: 33164642]
53. Zhang E, LeQuesne E, Fichtel K, Ginsberg D, Frankle W. In-patient psychiatry management of COVID-19: Rates of asymptomatic infection and on-unit transmission. BJPsych Open 2020 Sep 01;6(5):e99 [FREE Full text] [doi: 10.1192/bjo.2020.86] [Medline: 32867874]
54. Bäuerle A, Graf J, Jansen C, Musche V, Schweda A, Hetkamp M, et al. E-mental health mindfulness-based and skills-based 'CoPE It' intervention to reduce psychological distress in times of COVID-19: Study protocol for a bicentre longitudinal study. BMJ Open 2020 Aug 13;10(8):e039646 [FREE Full text] [doi: 10.1136/bmjopen-2020-039646] [Medline: 32792455]
55. Dursun OB, Turan B, Pakyürek M, Tekin A. Integrating telepsychiatric services into the conventional systems for psychiatric support to health care workers and their children during COVID-19 pandemics: Results from a national experience. Telemed J E Health 2021 Mar;27(3):269-275. [doi: 10.1089/tmj.2020.0237] [Medline: 32821025]
56. El Hayek S, Cheaito M, Nofal M, Abdelrahman D, Adra A, Al Shamli S, et al. Geriatric mental health and COVID-19: An eye-opener to the situation of the Arab countries in the Middle East and North Africa region. Am J Geriatr Psychiatry 2020 Oct;28(10):1058-1069 [FREE Full text] [doi: 10.1016/j.jagp.2020.05.009] [Medline: 32425474]

57. Feijt M, de Kort Y, Bongers I, Bierbooms J, Westerink J, IJsselsteijn W. Mental health care goes online: Practitioners' experiences of providing mental health care during the COVID-19 pandemic. Cyberpsychol Behav Soc Netw 2020 Dec;23(12):860-864. [doi: 10.1089/cyber.2020.0370] [Medline: 32815742]
58. Haxhihamza K, Arsova S, Bajraktarov S, Kalpak G, Stefanovski B, Novotni A, et al. Patient satisfaction with use of telemedicine in university clinic of psychiatry: Skopje, North Macedonia during COVID-19 pandemic. Telemed J E Health 2021 Apr;27(4):464-467. [doi: 10.1089/tmj.2020.0256] [Medline: 32809916]
59. Looi J, Allison S, Bastiampillai T, Pring W. Private practice metropolitan telepsychiatry in larger Australian states during the COVID-19 pandemic: An analysis of the first 2 months of new MBS telehealth item psychiatrist services. Australas Psychiatry 2020 Dec;28(6):644-648. [doi: 10.1177/1039856220961906] [Medline: 32997521]
60. Olwill C, Mc Nally D, Douglas L. Psychiatrist experience of remote consultations by telephone in an outpatient psychiatric department during the COVID-19 pandemic. Ir J Psychol Med 2021 Jun;38(2):132-139 [FREE Full text] [doi: 10.1017/ipm.2020.51] [Medline: 32438945]
61. Patel S, Gannon A, Cryan M, McDonnell N, Rafiq S, Adamis D, et al. Ordinary care in extraordinary times. Ir J Psychol Med 2020 Sep 11:1-5 [FREE Full text] [doi: 10.1017/ipm.2020.103] [Medline: 32912356]
62. Roncero C, García-Ullán L, de la Iglesia-Larrad JI, Martín C, Andrés P, Ojeda A, et al. The response of the mental health network of the Salamanca area to the COVID-19 pandemic: The role of the telemedicine. Psychiatry Res 2020 Sep;291:113252 [FREE Full text] [doi: 10.1016/j.psychres.2020.113252] [Medline: 32623263]
63. Conrad RC, Baum ML, Shah SB, Levy-Carrick NC, Biswas J, Schmelzer NA, et al. Duties toward patients with psychiatric illness. Hastings Cent Rep 2020 May;50(3):67-69 [FREE Full text] [doi: 10.1002/hast.1139] [Medline: 32596900]
64. Falala-Séchet C, Antoine L, Thiriez I. Owlie, un chatbot de soutien psychologique: Pourquoi, pour qui ? Inf Psychiatr 2020;96(8):659-666. [doi: 10.1684/ipe.2020.2168]
65. Genoud-Prachex T, Perrenoud M, Brioschi GA, Moser C, Bieler-Aeschlimann M, Rouaud O. Télé-neuropsychologie, patients âgés et Covid-19. Rev Neuropsychol 2020;12(2):178-180. [doi: 10.1684/nrp.2020.0564]
66. Goldin D, Maltseva T, Scaccianoce M, Brenes F. Cultural and practical implications for psychiatric telehealth services: A response to COVID-19. J Transcult Nurs 2021 Mar;32(2):186-190. [doi: 10.1177/1043659620973069] [Medline: 33150851]
67. Ibrahim FA, Pahuja E, Dinakaran D, Manjunatha N, Kumar CN, Math SB. The future of telepsychiatry in India. Indian J Psychol Med 2020 Oct;42(5 Suppl):112S-117S [FREE Full text] [doi: 10.1177/0253717620959255] [Medline: 33354056]
68. Jayarajan D, Sivakumar T, Torous JB, Thirthalli J. Telerehabilitation in psychiatry. Indian J Psychol Med 2020 Oct;42(5 Suppl):57S-62S [FREE Full text] [doi: 10.1177/0253717620963202] [Medline: 33354066]
69. Looi JC, Pring W. To tele- or not to telehealth? Ongoing COVID-19 challenges for private psychiatry in Australia. Australas Psychiatry 2020 Oct;28(5):511-513 [FREE Full text] [doi: 10.1177/1039856220950081] [Medline: 32847378]
70. Mohan A, Ambekar A. Telepsychiatry and addiction treatment. Indian J Psychol Med 2020 Oct;42(5 Suppl):52S-56S [FREE Full text] [doi: 10.1177/0253717620958169] [Medline: 33354065]
71. Cave J, Crews M. Rehabilitation during a pandemic: Psychiatrists as first responders? J Psychosoc Rehabil Ment Health 2020 Oct 22:1-2 [FREE Full text] [doi: 10.1007/s40737-020-00200-w] [Medline: 33110740]
72. Wasser T, Hauser L, Kapoor R. The management of COVID-19 in forensic psychiatric institutions. Psychiatr Serv 2020 Oct 01;71(10):1088-1090. [doi: 10.1176/appi.ps.202000584] [Medline: 32998660]
73. Gardner J, Plaven B, Yellowlees P, Shore J. Remote telepsychiatry workforce: A solution to psychiatry's workforce issues. Curr Psychiatry Rep 2020 Jan 27;22(2):8. [doi: 10.1007/s11920-020-1128-7] [Medline: 31989417]
74. Hilty DM, Sunderji N, Suo S, Chan S, McCarron RM. Telepsychiatry and other technologies for integrated care: Evidence base, best practice models and competencies. Int Rev Psychiatry 2018 Dec;30(6):292-309. [doi: 10.1080/09540261.2019.1571483] [Medline: 30821540]
75. Shore JH, Yellowlees P, Caudill R, Johnston B, Turvey C, Mishkind M, et al. Best practices in videoconferencing-based telemental health April 2018. Telemed J E Health 2018 Nov;24(11):827-832. [doi: 10.1089/tmj.2018.0237] [Medline: 30358514]
76. Yellowlees P, Richard Chan S, Burke Parish M. The hybrid doctor-patient relationship in the age of technology - Telepsychiatry consultations and the use of virtual space. Int Rev Psychiatry 2015;27(6):476-489. [doi: 10.3109/09540261.2015.1082987] [Medline: 26493089]
77. Améliorer la Santé Mentale de la Population: Vers Une Stratégie sur la Santé Mentale Pour l'Union Européenne. Brussels, Belgium: European Commission; 2005 Oct 14. URL: https://ec.europa.eu/health/archive/ph_determinants/life_style/mental/green_paper/mental_gp_fr.pdf [accessed 2022-01-26]
78. Kocsis BJ, Yellowlees P. Telepsychotherapy and the therapeutic relationship: Principles, advantages, and case examples. Telemed J E Health 2018 May;24(5):329-334. [doi: 10.1089/tmj.2017.0088] [Medline: 28836902]
79. Zhang M, Smith HE. Digital tools to ameliorate psychological symptoms associated with COVID-19: Scoping review. J Med Internet Res 2020 Aug 21;22(8):e19706 [FREE Full text] [doi: 10.2196/19706] [Medline: 32721922]
80. Shore J. Best practices in tele-teaming: Managing virtual teams in the delivery of care in telepsychiatry. Curr Psychiatry Rep 2019 Jul 08;21(8):77. [doi: 10.1007/s11920-019-1052-x] [Medline: 31286277]

81. Hilty DM, Rabinowitz T, McCarron RM, Katzelnick DJ, Chang T, Bauer AM, et al. An update on telepsychiatry and how it can leverage collaborative, stepped, and integrated services to primary care. Psychosomatics 2018;59(3):227-250. [doi: 10.1016/j.psym.2017.12.005] [Medline: 29544663]

Abbreviations

ARN: French National Research Agency
PHRC: French National Hospital Program for Clinical Research
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Kushniruk; submitted 11.05.21; peer-reviewed by S Bonneton, P Yellowlees; comments to author 14.06.21; revised version received 06.07.21; accepted 09.10.21; published 04.02.22.

Please cite as:
Kane H, Gourret Baumgart J, El-Hage W, Deloyer J, Maes C, Lebas MC, Marazziti D, Thome J, Fond-Harmant L, Denis F
Opportunities and Challenges for Professionals in Psychiatry and Mental Health Care Using Digital Technologies During the COVID-19 Pandemic: Systematic Review
JMIR Hum Factors 2022;9(1):e30359
URL: https://humanfactors.jmir.org/2022/1/e30359
doi:10.2196/30359
PMID:34736224

Original Paper

The Use of Telehealth for Psychological Counselling of Vulnerable Adult Patients With Rheumatic Diseases or Diabetes: Explorative Study Inspired by Participatory Design

Mette Juel Rothmann[1,2,3,4], MHS, PhD; Julie Drotner Mouritsen[1,2], MSc; Nanna Skov Ladefoged[1,2], MSc; Marie Nedergaard Jeppesen[2], MHH; Anna Sofie Lillevang[1,2], MCN; Helle Laustrup[1,4], MD, PhD; Torkell Ellingsen[1,4], MD, PhD

[1]Rheumatology Research Unit, Odense University Hospital, Odense, Denmark

[2]Steno Diabetes Center Odense, Odense University Hospital, Odense, Denmark

[3]Centre for Innovative Medical Technology, Odense University Hospital, Odense, Denmark

[4]Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Corresponding Author:
Mette Juel Rothmann, MHS, PhD
Steno Diabetes Center Odense
Odense University Hospital
J.B. Winsløws Vej 4
Odense, 5000
Denmark
Phone: 45 22346388
Email: mette.rothmann@rsyd.dk

Abstract

Background: Video consultation is increasingly used in different health care settings to reach patients. However, little is known about telehealth in psychological counselling for vulnerable patients with somatic and chronic conditions such as rheumatoid arthritis and diabetes.

Objective: This study aimed to develop and pilot test a telepsychology module for inclusion in the app My Hospital (Mit Sygehus) to provide remote psychological counselling to vulnerable adults with either rheumatic diseases or diabetes.

Methods: With inspiration from participatory design, the content of the telepsychology module was developed through user involvement and evaluated by individual interviews with patients and psychologists as well as questionnaires.

Results: We developed a module with our patient partners that targeted patients with rheumatic diseases and diabetes in relation to the psychological challenges of living with chronic diseases. The module included information, tools, exercises, and videoconferencing. In total, 16 patients and 3 psychologists participated in the pilot test. Psychological counselling was described by 4 themes: “The good relation despite physical distance,” “The comfort of being at home,” “The pros of saving time on transport and energy,” and “A therapeutic alliance at a distance.”

Conclusions: Psychological counselling in relation to somatic care can be provided by videoconferencing supported by web-based or mobile delivery of tailored information, tools, and exercises without compromising on the quality of care. To ensure a good alliance between the patient and psychologist, a first face-to-face meeting is important. The home location provided patients with a safe environment and increased accessibility and reduced travel time to the hospital.

(JMIR Hum Factors 2022;9(1):e30829) doi:10.2196/30829

KEYWORDS

telehealth; videoconferencing; app; co-production; co-creation; psychologist; psychology; rheumatic diseases; diabetes; mobile phone

Introduction

Telehealth in rheumatic and diabetes care is increasingly used to reach patients in rural areas as well as to maximize patient care in terms of reduced travel time to hospital sites and the convenience of consulting with clinicians from the patients' own homes [1-4]. Lately, the COVID-19 situation has changed the way health care services are being delivered, with an increase in the use of telehealth [5,6]. Hence, there is currently a momentum for telehealth globally [6], which is important to build on, as the uptake of telehealth (eg, videoconferencing) in clinical practice has been slower than expected because of factors such as immature technology or unwillingness of health care professionals to adopt the new service [7,8]. In general, research has shown that patients are satisfied with psychotherapy provided through telehealth [9-11]. Furthermore, telehealth and telemental health have been shown to be as effective as in-person care [12,13]. Thus, several studies have shown videoconferencing to be safe and to have the same (or better) clinical effect as traditional consultations. Among others, these studies included patients with diabetes, heart failure, cancer, depression, and posttraumatic stress disorder [8,10,12,14]. In addition, a review on telepsychology has shown that video or phone sessions are effective for the treatment of conditions such as depression and anxiety [15]. Despite the growing empirical support for telemental health and telepsychology with regard to efficacy and patient satisfaction, literature on telepsychology for patients with somatic chronic conditions, such as rheumatoid arthritis and diabetes, is sparse. The term telepsychology is defined as "the provision of psychological services using telecommunication technologies" [16].

It is well known that patients with rheumatic diseases, such as rheumatoid arthritis, face challenges because of symptoms including joint or muscle pain and fatigue [17] and that many aspects of daily life in patients with diabetes are also significantly impacted because of psychological challenges, depression, and anxiety [18,19]. These challenges as well as the risk of depression, anxiety, and disease-specific distress can result in significant health consequences because of reduced coping strategies and self-management. Hence, psychological support and treatment are important in the care of patients with rheumatic diseases and diabetes [20,21]. However, there might be barriers to seeking help, which include comorbidities, geographical and time constraints, limited personal resources, and individual concerns. Even though Denmark is a small country, distance to the hospital is an important factor for inequality in the availability of health care services. In the long run, distance will also affect the individual patient's ability to receive treatment because of the development in the Danish health care system where many treatment options (eg, disease-specific psychological counselling) are grouped at national university hospitals (6 in Denmark). Also, finances can be a barrier if patients are referred for psychological care (self-payment) by their general practitioner. Finally, it is inevitable that health care systems must change the manner in which health services are delivered because of demographic changes in patient populations, that is, increased aging population and burden of diseases, and must ensure equal access to health care. Telepsychology might be a way to address these barriers. A literature search has revealed sparse knowledge on telepsychology, telehealth, and psychological counselling for patients with rheumatic diseases as part of their somatic treatment. However, web-based psychoeducational intervention for patients with fibromyalgia syndrome has been shown to be effective on psychological variables [22]. Similar results were seen for a web-based program for youth with juvenile idiopathic arthritis [23]. Finally, a web-based, tailored cognitive behavioral intervention for patients with rheumatoid arthritis and a psychological risk profile has shown a positive effect on psychological outcomes [24]. In the field of diabetes, studies have demonstrated the effect of a web-based self-help tool in improving patients' psychological well-being [25]. Group-based sessions with a focus on fear of hypoglycemia delivered remotely through telemedicine to parents of young children with type 1 diabetes (1-6 years of age) [26] and an mHealth service with automated interactive voice response that monitors patients' self-management, provides immediate problem-tailored support, and connects to clinicians and the appropriate family members for feedback have also shown to be effective [27]. Furthermore, web-based psychotherapy programs have been shown to be effective in reducing depressive symptoms in patients with both type 1 and 2 diabetes [28]. There is a lack of knowledge about the use and acceptability of telepsychology compared to usual practice (same-room sessions). However, existing literature underpins that web-based psychotherapy may be effective in providing psychological counselling remotely. One such tool for telepsychology is the Region of Southern Denmark's app My Hospital (Mit Sygehus). Launched in 2014, My Hospital provides relevant information to patients and their relatives. It consists of modules for each department in the region's hospitals. My Hospital is easily available and free of charge. My Hospital is available in the App Store and Google Play as well as in a web version [29]. In general, the tools in the app are available without the need for individual user login. However, a personal login is needed to use features of the app where personal information and data are collected, such as videoconferencing, to comply with data protection regulation (the personal login for the app is secured and encrypted).

Hence, the purpose of this study was to develop and pilot test a telepsychology module for inclusion in My Hospital to provide remote psychological counselling to vulnerable adult patients with either rheumatic diseases or diabetes.

Methods

This study was inspired by participatory design where the idea is to engage users to innovate and develop technologies in collaboration with developers [30]. In a traditional participatory design project, the users would be engaged from the beginning and throughout the study to ensure that the technology meets the end user's needs [31]. In our case, we did not follow the participatory design methodology strictly, but engaged a former patient to participate in the development of the telepsychology module along with a colleague specialized in communication, information technology (IT) specialists, clinicians, and researchers. After the development phase, the patients referred to psychological counselling were invited to participate in the

project and pilot test the module, including videoconferencing, and provide input for adjustments or changes.

This study was inspired from hermeneutics, where the perspective has been to understand the participants' lived experiences in relation to living with rheumatologic diseases and diabetes to develop a service that meets the patients' needs [32].

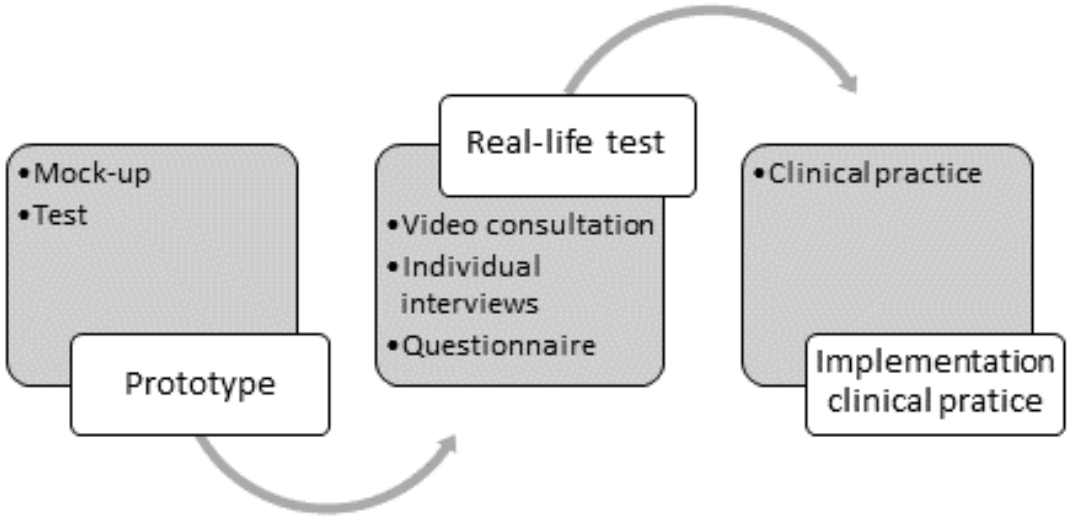
Sample and Context

The participants were treated, followed, and recruited at the Department of Rheumatology and Steno Diabetes Center Odense, Odense University Hospital, Denmark. Patients were referred by nurses or doctors at the outpatient clinic to the department's psychologists because of psychological challenges arising from living with chronic rheumatic disease or diabetes. The sampling was purposive and patients aged ≥18 years with rheumatic diseases or diabetes were considered eligible; however, the psychologists assessed the burden of disease. Thus, patients excluded were those who did not have any device at home (computer or tablet) and those who were assessed by the psychologist to be too sick for inclusion in the study, such as patients with very severe psychological problems. Furthermore, 3 psychologists participated in the study. The psychologists were master-level counsellors: 1 was newly qualified, 1 had a few years of experience, and 1 had more than 5 years of clinical experience.

Data Collection

Data collection was divided into 2 processes (Figure 1) and performed between October 2018 and December 2019.

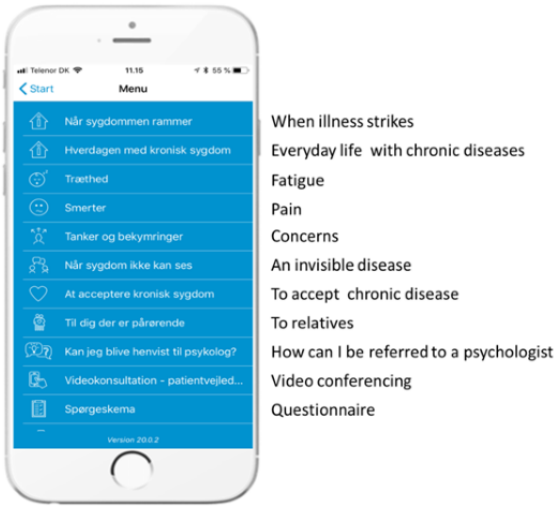
Figure 1. Data collection.



Development of Prototype

First, a mock-up of the module including information and tools was developed based on clinical and specialist knowledge from the psychologist (unblended when published). The mock-up was discussed and re-designed in collaboration with a patient diagnosed with rheumatoid arthritis and a professional communicator specialized in patient involvement. The information and tools in the module were adjusted in an iterative process until they were accepted by all. Data were collected through track-change, notes, and meetings. Based on the feedback and changes, a prototype of the module was developed and included in My Hospital. The module included information on aspects such as everyday life with chronic disease, concerns, and acceptance, a worksheet with the cognitive behavioral model [33], and a description of mindfulness-based exercises. Thus, these could be used as basis for homework between the therapy sessions. Figure 2 shows a screenshot of My Hospital. The next step was test sessions, as videoconferencing has not been used for psychological counselling. My Hospital was updated to include videoconferencing. First, the technology for videoconferencing was tested, and login instructions were developed together with the patient, clinician, and IT specialist. The app and videoconference were integrated (partly) in the electronic patient record, allowing easy access for clinicians and integration to the booking system. The second author (JDM) conducted test sessions together with the patient to facilitate mutual learning throughout counselling. Psychological counselling was mainly based on cognitive behavioral therapy [33] and acceptance and commitment therapy [34]. Counselling consisted of conversation, exercises, and homework. The test sessions had a special focus on communication and the therapeutic alliance in relation to the counselling. Data on shared reflection and experiences with the technology were collected through notes.

Figure 2. My Hospital.



Real-life Test of the Service—the Intervention

Three psychologists from the departments were involved in the service, and 16 patients were invited to participate in the real-life test, in which usual counselling (individual planned sessions including 2-8 sessions) was converted into videoconferencing supported by the module in My Hospital and included information, tools, and exercises and homework between sessions. However, the first session was performed face-to-face at the hospital to ensure good relation as an important starting point for counselling. Furthermore, the psychologist was able to assess whether videoconferencing would be relevant and safe considering the severity of psychological problems.

The patient was given written instructions on how to access the app or web platform My Hospital and how to use My Hospital for videoconferencing as well as relevant contact information (phone numbers for the psychologists and IT support). The patient was verbally informed about the option of videoconferencing as an alternative to same-room counselling, necessary equipment (tablet or computer with a microphone and a camera, a well-functioning internet connection), a plan B in case of technical issues or dissatisfaction with the videoconferencing format (telephone counselling or same-room counselling), and the option of trying out videoconferencing with IT support before the first videoconference with the psychologist.

Results from the first phase showed that necessary equipment included a computer or a tablet with a camera, a microphone, and a strong internet connection. Smartphones were not recommended because of the small screen size and because they are hand-held devices. The study was based on “bring your own device” to minimize cost. Patients were automatically notified 15 minutes before a preplanned videoconference through a notification from My Hospital. This allowed the patient to “sign in” to the video platform so that the psychologist could start the videoconference. Data from the testing phase were collected through individual, semi-structured interviews with patients, and psychologists. MJR and JDM carried out the interviews. All interviews with patients were conducted by telephone or in the same room. The interviews varied between 30 and 60 minutes. All interviews were audio recorded. In addition, a short questionnaire was automatically sent to the patient by My Hospital after each videoconference. Furthermore, statistics from My Hospital showing the “click-rate” on different topics and indicating the topics that the patients read the most were collected. Finally, psychologists were asked to fill in log books on experiences with videoconference.

Analysis

The data from the development phase were used to adjust and re-design the tool as well as to create preliminary guidelines for the service using a PDSA (Plan-Do-Study-Act) cycle approach [35]. The qualitative data from the second phase (real-life test) were analyzed by Braun and Clark’s text condensation [36]. First, we (MJR, MNJ) captured an overall impression of the data and extracted a preliminary set of main themes. Second, data were divided into meaningful topics relevant to the research question. Then, the topics were condensed and coded. Finally, the findings were synthesized, which involved a shift from condensation to descriptions and categories.

Ethics

According to Danish law, qualitative studies do not require approval from scientific ethics committees. However, this study was approved by the Danish Data Protection Agency (18/51158). The participants received both oral and written information before signing informed consent in compliance with the Helsinki

Declaration. Data were stored and secure in a logged SharePoint, Region of Southern Denmark.

Results

Results from the real-life pilot test are presented below. In total, 16 people agreed to participate in this study. Characteristics of the participants are shown in Table 1.

Table 1. Characteristics of participants (n=16).

Characteristic	Value
Gender, n	
Men	3
Women	13
Age (years), range	25-59
Diagnosis, n	
Rheumatologic diseases	13
Diabetes	3
Number of sessions	
Mean	4.3
Range	2-10
Types of session, range	
In-person session	1-3
Video session	1-6
Telephone session	0-3
Distance to hospital, km	
Mean	53
Range	4-130

A total of 45 videoconferences were conducted. All counselling sessions were of a shorter duration, varying from 1 to 10 consultations (including face-to-face and telephone consultations). Of the 16 patients, 6 (38%) agreed to participate in an individual interview. The remaining 10/16 (62%) patients were not able to participate because of their psychological state. Furthermore, individual interviews were conducted by 2 psychologists, and questionnaires were answered by patients after each videoconference.

Results From Individual Interviews With Patients

The thematic analysis revealed the following 3 themes: “The good relation despite physical distance,” “The comfort of being at home,” and “The pros of saving time on transport and energy.”

The Good Relation Despite Physical Distance

In general, all participants had a good relationship with the psychologist despite the use of videoconferencing instead of same-room counselling. The foundation was established in a first same-room meeting, which all the participants described as very important to establish a trusting and safe environment. One woman said, “Initially, you do not know who you are going to meet, so it is nice to meet face-to-face” (Patient #4). However, one participant stressed that even though the relation was different from that in a same-room meeting, it was still good enough to ensure valuable counselling.

> *It’s different (video conferencing)...although I would like to give in…it’s just not the same as when we sit in the same room.* [Patient #2]

On the other hand, some (3/6) of the other participants stated that videoconferencing was as good as same-room meetings. Thus, the majority (4/6) of participants did not see videoconferencing as a barrier for the topics and issues to be discussed. One major reason for that was the professional skills of the psychologist in ensuring high-quality counselling through videoconferencing.

> *The psychologist is really good and she does well on video. She is present and the relationship is almost as good as when we meet face-to-face (in the same-room).* [Patient #2]

The participants felt that the psychologist not only had professional skills but also had special skills in relation to the use of technology and videoconferencing, that is, making eye contact, using appropriate tone of voice, pausing, and using tools.

Although the experience of videoconferencing use was positive, technical problems were described as an important barrier. The majority (5/6) of participants experienced different types of technical problems, all of which affected the session. However, the participants managed to continue the session by video, and they highlighted the importance of a “plan B” if there were

problems such as knowing how to reconnect to video or switch to a telephone call.

The Comfort of Being at Home

In general, the participants had a very positive experience of the "home" sessions, as their home provides a safe environment. As 1 participant said, "I could cry...I had a secure environment, it's my home after all" (Patient #4). However, the participants stressed that there are certain challenges when therapy is provided at home. Sessions at home were new to everyone in the household and this required special attention. The participants had to find a suitable room for the conversation, and some (2/6) of the participants experienced lack of respect from family members. One participant said, "they just opened the door and came in…my husband…Oh, I just wanted to pick something up" (Patient #2). Lack of privacy was described by some (2/6) of the participants as something they had to talk about with their family members. On the other hand, one of the participants had her husband join the session. In general, the possible barriers at home were minor compared to the benefits of staying at home to "avoid transport and fatigue" (Patient #5).

The Pros of Saving Time on Transport and Energy

All participants described the benefit of not having to go to the hospital for counselling. Even for patients who live near the hospital, it could be stressful and take a lot of energy not only because of driving but also because finding a parking space can sometimes be impossible. Patients with rheumatic diseases (6/6), in particular, found it beneficial to stay at home, as their conditions were often associated with severe fatigue. As 1 patient said, "... when you do not have any energy left it doesn't matter if you have to drive 1 or 50 kilometers" (Patient #4). The energy saved by staying at home was valued as important, and it enabled patients to participate in sessions and not cancel even on "bad days," as the energy saved could be spent on counselling instead. One participant said, "So even though it had been a day with pain and lack of energy, I didn't have to take the car, I just had to open my computer" (Patient #4). All participants described the benefit of saving time on transportation as saving energy, which enabled them to "use your energy on the right things" (Patient #6).

Results From Individual Interviews With Psychologists

The interviews with psychologists revealed another meaningful theme: "A therapeutic alliance at a distance.

A Therapeutic Alliance at a Distance

In general, the psychologists were a little skeptical about the use of videoconferencing, and they questioned whether it was good enough and whether the therapeutic alliance could be built and maintained. They were worried about technical failure, as it would affect the alliance and leave patients on their own. On the other hand, the psychologists described how patients might benefit from the service, as many patients deselect same-room counselling because they already have too many appointments at the hospital and lack the personal resources to participate.

> *But there is also a unique opportunity to reach out to those who do not have the personal resources as it is takes too much energy and might have anxiety as well.* [Psychologist #1]

Thus, the benefits of videoconferencing compensated for the risk of technical problems. With time, the psychologists experienced that the therapeutic alliance was the same as that in same-room counselling. However, to ensure a good patient-psychologist alliance, both psychologists stated that the first session should always be face-to-face in the same room. Furthermore, they described how videoconferencing actually provided extra information, as the psychologist met the patient in their home and got a view of the patient's daily life.

> *…it also affects the alliance as I can see a part of their life that I do not otherwise have access to, namely their home.* [Psychologist #2]

The benefits for patients were evident for the psychologist, as the home provided a safe environment for the patient. One of the psychologists said, "... I actually do not think there is a big difference (in the contact and alliance same-room versus video conferencing)…" [Psychologist #2].

However, it took courage to try videoconferencing because psychologists have to move away from the well-known consultation room to video. They described a feeling of losing control because of the unknown. In this new "room," technology played a significant role because the risk of failure and technical breakdowns led to concerns among the psychologists, as it could prevent delivery of good, high-quality service. Especially in complex cases, the psychologist needed a sense of safety and in these cases, they were more concerned about the technology and distance to the patient.

> *...if I picked up something like a trauma, I was very aware of getting the patient to a same-room counselling next time.* [Psychologist #1]

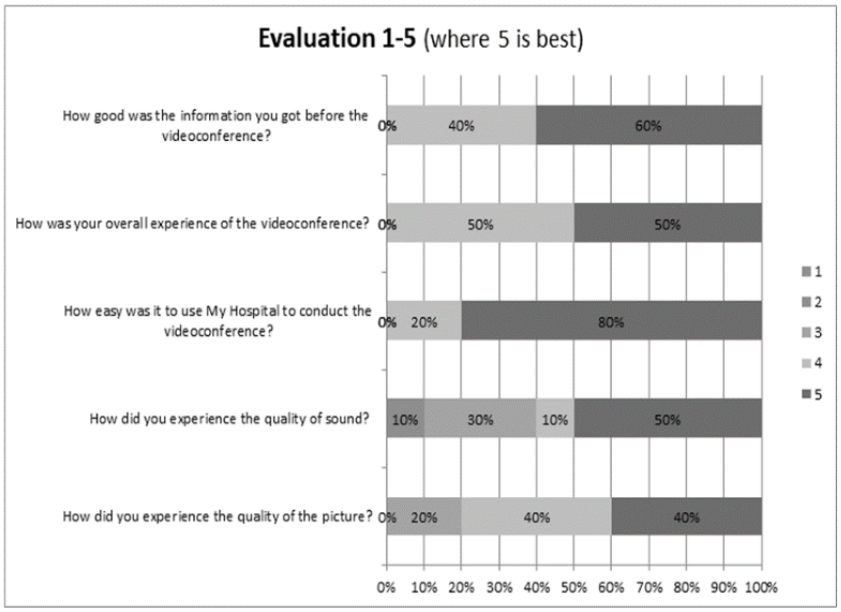
Despite the concerns about the technology, the psychologists found that the videoconferences combined with the information and tools in the app were a good alternative to face-to-face sessions even though the technology might break down. Becoming familiar with the service was described as important, and the psychologists described how they had to develop new communications skills for videoconferencing to facilitate a good therapeutic alliance, for example, a plan B if the video failed.

> *...the patients also adapt fairly quickly (red. to video conferencing) when it has small delay in the sound or picture… you might interrupt each other…but it was okay as it did not impact on the conversation or the therapeutic alliance.* [Psychologist #2]

Results From the Questionnaires

A brief questionnaire was released in My Hospital after each videoconference with the purpose of obtaining the patient's immediate assessment of the counselling. A total of 16 patients completed the questionnaire. Data showed that the most common reason for choosing a virtual session again was "saving time on transport" followed by "Physical challenges e.g. parking, stairs." The questionnaire also evaluated the information, accessibility, and quality of the virtual session. The results are shown in Figure 3.

Figure 3. Evaluation of the service (n=16).



Overall, videoconferencing was associated with a high degree of satisfaction with respect to information about the videoconferencing, the overall experience, the use of My Hospital, and the quality of sound and picture. However, data showed that there is room for improvement, as 6/16 (38%) participants rated sound quality as average or poor and 3/16 (19%) participants rated the quality of the picture as average.

Data From the App

The telepsychology module including information and tools in the form of text, videos, audio files, pictures, exercises, and videoconferencing was included in My Hospital. Once patients had been assigned to a treatment pathway in My Hospital, they could see an overview of topics related to the psychosocial challenges that can arise from living with a chronic disease. Data from My Hospital showed that topics such as “fatigue,” “pain,” “when illness strikes,” and “when illness is invisible” had the most views.

Discussion

Principal Findings

This study aimed at developing and pilot testing a telepsychology module for inclusion in My Hospital to provide remote psychological counselling to vulnerable adult patients with either rheumatic diseases or diabetes and psychological problems. The findings revealed that telepsychology and psychological counselling in relation to somatic care can be provided by videoconferencing and supported by tailored information, tools, and exercises.

To meet the needs of patients, we co-created a module in My Hospital together with a patient, a psychologist, a staff member specialized in patient communication, and IT experts to ensure inclusion of relevant knowledge from all participants [37]. The module contained information on aspects including everyday life with chronic disease, concerns, and acceptance, a worksheet with the cognitive behavioral model [33], and a description of acceptance and commitment therapy–based mindfulness exercises [34]. In line with previous studies using web-based tools [22-25,28,38,39] we found the module to be feasible and accepted by the patients. However, the module did differ from that in previous studies in that our module was designed to support the psychological counselling provided by videoconferencing. Hence, feedback on worksheets and exercises was given by the psychologist during counselling and not as interactive self-help [25], email messaging service [24,39], or phone calls [23]. Our study revealed that the patients found psychological counselling provided by videoconferencing to be a good alternative to same-room counselling at the hospital. This is supported by the findings of prior research evaluating telephone versus videoconferencing [40] and videoconferencing versus same-room–delivered therapy [41]. However, our study sample was relatively small, and it cannot be concluded whether videoconferencing is appropriate for all patients. Thus, the psychologists assessed some of the patients to be too vulnerable and sick for videoconferencing for reasons such as trauma. There is a need for evidence that can help us define specific populations who may most benefit from telepsychology, including videoconferencing and web-based counselling, and who may not. However, there is no evidence that counselling provided by videoconferencing can be harmful, and in general, it seems that patients benefit from the use of videoconferencing and web-based tools [2,42,43]. Still, existing literature shows that clinicians’ willingness to use and acceptance of telehealth are important for implementation [7]. In line with prior research findings on cost, travel time, and flexibility [26,44,45], our findings highlight the importance of being at home and saving time and energy, which can be of great importance to vulnerable patients. However, Matsumoto and Barton [2] point out that the

increasing use of telehealth might introduce inequality because of lack of access (high-speed internet and smartphones). On the other hand, one could argue that patients might not get access to counselling if clinicians do not adopt the new service because of barriers such as cost and travel time. To minimize some of the barriers in telehealth, it is also important that the patients are interested in the service, have adequate IT skills, and have access to the necessary equipment [42]. We found that the module included in My Hospital was easy to access and use from the user's own device (PC, tablet, or smartphone). Smartphones worked well for the module targeting information on the psychological impact of chronic diseases, but PCs or tablets were recommended for the videoconferencing to support a good relation and the psychological alliance, as they provide better picture, sound, and view of body language. Today, internet access and mobile technologies are an integrated part of everyday life for a large proportion of the population in industrialized countries, including Denmark. According to the most recent data from Statistics Denmark (2020), 95% of the Danish population has internet access [46], 90% have smartphones, 60% of families have tablets, 88% have laptops, and 36% have static PCs [47] in their homes. Thus, there might be a small proportion of the population without access to telehealth services. This could be managed by hospitals through the offering of tablets on loan to patients who need it. Moving of the therapy room to the home was well accepted, as the home provided a safe environment. However, this changed the setting at home, as the patients and their families were unfamiliar with videoconferencing and did not respect or think of the need for privacy during the sessions. Our findings stress the need for articulating issues of privacy in the family and to ensure that the patients are able to find a private space at home for the sessions. In addition, our findings revealed that it is important to have a clear-cut plan B if the technology or internet connection fails, as it could leave the patient vulnerable. In line with findings from studies on telemental health [11] and internet-based cognitive behavioral therapy [48], this study found that videoconferencing facilitated the therapeutic alliance as well as same-room counselling.

Strengths and Limitations

A limitation of this study is that it was a small-scale study. However, the aim of this study, like other small-scale qualitative studies, was to provide in-depth exploration of the phenomenon under investigation. Therefore, the intention of this study was to develop and investigate the potential of telepsychology, including web-based tools and videoconferencing, for psychological counselling of patients with either rheumatic diseases or diabetes. It has to be taken into account that only 16 patients participated in the test, and that only 6 (38%) of these accepted participation in an individual interview. However, it is important to acknowledge that this study included vulnerable patients with severe psychological challenges, so the low degree of participation in interviews was somewhat expected.

That said, we have provided rich descriptions of both the development of the module (eg, information, worksheets, exercises) and potential of videoconferencing. The analysis was conducted in collaboration with co-researchers to increase reliability. To warrant validity, quotes from the interviews were used to link to the participants' original statements. Hopefully, this will allow the readers to judge whether the findings of this study are transferable to their own contexts.

Conclusions

Telepsychology and psychological counselling in relation to somatic care can be provided by videoconferencing and supported by web-based or mobile-delivered tailored information and tools. However, to ensure a good patient-psychologist alliance, the first session should always be in the same room. With a well-established alliance, the patients' homes provided a safe environment and enabled them to conserve their energy that could be used on the right things, which is of great importance especially to patients with rheumatic diseases and severe fatigue. In patients with diabetes, the benefits were accessibility and reduced travel time to the hospital.

Acknowledgments

We wish to thank all the participants involved in the study and especially the patient representatives who contributed to the manuscript. We also wish to acknowledge Lea Bohn, Centre for Medical Technology (CIMT), Odense University Hospital, for assistance in reviewing the manuscript. This study was supported by the Odense University Hospital Innovation fund.

Conflicts of Interest

None declared.

References

1. Nelson AJ, Anderson ME. A systematic review exploring pre-COVID-19 telehealthcare models used in the management of patients with rheumatological disease. Rheumatol Adv Pract 2021;5(3):rkab073 [FREE Full text] [doi: 10.1093/rap/rkab073] [Medline: 34805737]
2. Matsumoto RA, Barton JL. Telerheumatology: before, during, and after a global pandemic. Curr Opin Rheumatol 2021 May 01;33(3):262-269. [doi: 10.1097/BOR.0000000000000790] [Medline: 33741808]

3. Timpel P, Oswald S, Schwarz PEH, Harst L. Mapping the evidence on the effectiveness of telemedicine interventions in diabetes, dyslipidemia, and hypertension: an umbrella review of systematic reviews and meta-analyses. J Med Internet Res 2020 Mar 18;22(3):e16791 [FREE Full text] [doi: 10.2196/16791] [Medline: 32186516]
4. Holt RI. The role of telehealth and diabetes. Diabet Med 2019 May;36(5):529-530. [doi: 10.1111/dme.13958] [Medline: 30980434]
5. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: Implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Jun;26(5):309-313 [FREE Full text] [doi: 10.1177/1357633X20916567] [Medline: 32196391]
6. Thomas EE, Haydon HM, Mehrotra A, Caffery LJ, Snoswell CL, Banbury A, et al. Building on the momentum: Sustaining telehealth beyond COVID-19. J Telemed Telecare 2020 Sep 26:1357633X20960638. [doi: 10.1177/1357633X20960638] [Medline: 32985380]
7. Wade VA, Eliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. Qual Health Res 2014 May;24(5):682-694. [doi: 10.1177/1049732314528809] [Medline: 24685708]
8. 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 Apr 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
9. Drago A, Winding T, Antypa N. Videoconferencing in psychiatry, a meta-analysis of assessment and treatment. Eur Psychiatry 2016 Aug;36:29-37. [doi: 10.1016/j.eurpsy.2016.03.007] [Medline: 27311105]
10. Ignatowicz A, Atherton H, Bernstein CJ, Bryce C, Court R, Sturt J, et al. Internet videoconferencing for patient-clinician consultations in long-term conditions: a review of reviews and applications in line with guidelines and recommendations. Digit Health 2019;5:2055207619845831 [FREE Full text] [doi: 10.1177/2055207619845831] [Medline: 31069105]
11. Jenkins-Guarnieri MA, Pruitt LD, Luxton DD, Johnson K. Patient perceptions of telemental health: systematic review of direct comparisons to in-person psychotherapeutic treatments. Telemed J E Health 2015 Aug;21(8):652-660. [doi: 10.1089/tmj.2014.0165] [Medline: 25885491]
12. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. Telemed J E Health 2013 Jun;19(6):444-454 [FREE Full text] [doi: 10.1089/tmj.2013.0075] [Medline: 23697504]
13. Norwood C, Moghaddam NG, Malins S, Sabin-Farrell R. Working alliance and outcome effectiveness in videoconferencing psychotherapy: A systematic review and noninferiority meta-analysis. Clin Psychol Psychother 2018 Nov;25(6):797-808. [doi: 10.1002/cpp.2315] [Medline: 30014606]
14. Armfield NR, Bradford M, Bradford NK. The clinical use of Skype--For which patients, with which problems and in which settings? A snapshot review of the literature. Int J Med Inform 2015 Oct;84(10):737-742. [doi: 10.1016/j.ijmedinf.2015.06.006] [Medline: 26183642]
15. Varker T, Brand RM, Ward J, Terhaag S, Phelps A. Efficacy of synchronous telepsychology interventions for people with anxiety, depression, posttraumatic stress disorder, and adjustment disorder: a rapid evidence assessment. Psychol Serv 2019 Nov;16(4):621-635. [doi: 10.1037/ser0000239] [Medline: 29809025]
16. Joint Task Force for the Development of Telepsychology Guidelines for Psychologists. Guidelines for the practice of telepsychology. Am Psychol 2013 Dec;68(9):791-800. [doi: 10.1037/a0035001] [Medline: 24341643]
17. Primdahl J, Hegelund A, Lorenzen AG, Loeppenthin K, Dures E, Appel Esbensen B. The experience of people with rheumatoid arthritis living with fatigue: a qualitative metasynthesis. BMJ Open 2019 Mar 20;9(3):e024338 [FREE Full text] [doi: 10.1136/bmjopen-2018-024338] [Medline: 30898808]
18. Jones A, Olsen MZ, Perrild HJ, Willaing I. The psychological impact of living with diabetes: descriptive findings from the DAWN2 study in Denmark. Prim Care Diabetes 2016 Feb;10(1):83-86. [doi: 10.1016/j.pcd.2015.03.008] [Medline: 25890864]
19. Roodenrijs NMT, van der Goes MC, Welsing PMJ, Tekstra J, Lafeber FPJG, Jacobs JWG, et al. Difficult-to-treat rheumatoid arthritis: contributing factors and burden of disease. Rheumatology (Oxford) 2021 Aug 02;60(8):3778-3788. [doi: 10.1093/rheumatology/keaa860] [Medline: 33331946]
20. American Diabetes Association. 5. Facilitating behavior change and well-being to improve health outcomes: Standards of Medical Care in Diabetes-2021. Diabetes Care 2021 Jan;44(Suppl 1):S53-S72. [doi: 10.2337/dc21-S005] [Medline: 33298416]
21. Roodenrijs NMT, Hamar A, Kedves M, Nagy G, van Laar JM, van der Heijde D, et al. Pharmacological and non-pharmacological therapeutic strategies in difficult-to-treat rheumatoid arthritis: a systematic literature review informing the EULAR recommendations for the management of difficult-to-treat rheumatoid arthritis. RMD Open 2021 Jan;7(1):e001512 [FREE Full text] [doi: 10.1136/rmdopen-2020-001512] [Medline: 33419871]
22. Conversano C, Poli A, Ciacchini R, Hitchcott P, Bazzichi L, Gemignani A. A psychoeducational intervention is a treatment for fibromyalgia syndrome. Clin Exp Rheumatol 2019;37 Suppl 116(1):98-104 [FREE Full text] [Medline: 30747101]
23. Stinson JN, McGrath PJ, Hodnett ED, Feldman BM, Duffy CM, Huber AM, et al. An internet-based self-management program with telephone support for adolescents with arthritis: a pilot randomized controlled trial. J Rheumatol 2010 Sep;37(9):1944-1952. [doi: 10.3899/jrheum.091327] [Medline: 20595280]

24. Ferwerda M, van Beugen S, van Middendorp H, Spillekom-van Koulil S, Donders ART, Visser H, et al. A tailored-guided internet-based cognitive-behavioral intervention for patients with rheumatoid arthritis as an adjunct to standard rheumatological care: results of a randomized controlled trial. Pain 2017 May;158(5):868-878. [doi: 10.1097/j.pain.0000000000000845] [Medline: 28106666]
25. Orman J, Clarke J, Whittle E, Anonuevo C, Proudfoot J. A diabetes-specific e-mental health tool: development, acceptability and outcomes of a feasibility study. Aust Fam Physician 2016 Aug;45(8):600-605 [FREE Full text] [Medline: 27610452]
26. Marker AM, Monzon AD, Nelson E, Clements MA, Patton SR. An intervention to reduce hypoglycemia fear in parents of young kids with type 1 diabetes through video-based telemedicine (REDCHiP): trial design, feasibility, and acceptability. Diabetes Technol Ther 2020 Jan;22(1):25-33 [FREE Full text] [doi: 10.1089/dia.2019.0244] [Medline: 31517512]
27. Aikens JE, Rosland A, Piette JD. Improvements in illness self-management and psychological distress associated with telemonitoring support for adults with diabetes. Prim Care Diabetes 2015 Apr;9(2):127-134 [FREE Full text] [doi: 10.1016/j.pcd.2014.06.003] [Medline: 25065270]
28. van Bastelaar KMP, Pouwer F, Cuijpers P, Riper H, Snoek FJ. Web-based depression treatment for type 1 and type 2 diabetic patients: a randomized, controlled trial. Diabetes Care 2011 Feb;34(2):320-325 [FREE Full text] [doi: 10.2337/dc10-1248] [Medline: 21216855]
29. Mit Sygehus. Google Play. URL: https://play.google.com/store/apps/details?id=dk.medware.mitforloeb&hl=da&gl=US [accessed 2022-03-02]
30. Clemensen J, Rothmann MJ, Smith AC, Caffery LJ, Danbjorg DB. Participatory design methods in telemedicine research. J Telemed Telecare 2017 Oct;23(9):780-785. [doi: 10.1177/1357633X16686747] [Medline: 28027678]
31. Simonsen JR. Routledge International Handbook of Participatory Design. New York: Routledge; 2012:295.
32. Hans Georg Gadamer. Truth and Method. UK: Bloomsbury Academic; 2013.
33. Beck JS. Cognitive Behavior Therapy, Third Edition: Basics and Beyond. New York: Guilford Publications; 2020.
34. Luoma JB, Hayes SC, Walser RD. Learning ACT: An Acceptance and Commitment Therapy Skills Training Manual for Therapists. In: Behavioral and Cognitive Psychotherapy. Oakland, CA: New Harbinger Publications; 2007:352.
35. Christoff P. Running PDSA cycles. Curr Probl Pediatr Adolesc Health Care 2018 Aug;48(8):198-201. [doi: 10.1016/j.cppeds.2018.08.006] [Medline: 30270135]
36. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
37. Kushniruk A, Nøhr C. Participatory design, user involvement and health IT evaluation. Stud Health Technol Inform 2016;222:139-151. [Medline: 27198099]
38. Ammerlaan J, van Os-Medendorp H, Scholtus L, de Vos A, Zwier M, Bijlsma H, et al. Feasibility of an online and a face-to-face version of a self-management program for young adults with a rheumatic disease: experiences of young adults and peer leaders. Pediatr Rheumatol Online J 2014 Mar 25;12:10 [FREE Full text] [doi: 10.1186/1546-0096-12-10] [Medline: 24666817]
39. Vallejo MA, Ortega J, Rivera J, Comeche MI, Vallejo-Slocker L. Internet versus face-to-face group cognitive-behavioral therapy for fibromyalgia: a randomized control trial. J Psychiatr Res 2015 Sep;68:106-113. [doi: 10.1016/j.jpsychires.2015.06.006] [Medline: 26228408]
40. Mochari-Greenberger H, Vue L, Luka A, Peters A, Pande RL. A tele-behavioral health intervention to reduce depression, anxiety, and stress and improve diabetes self-management. Telemed J E Health 2016 Aug;22(8):624-630. [doi: 10.1089/tmj.2015.0231] [Medline: 26954880]
41. Egede LE, Walker RJ, Payne EH, Knapp RG, Acierno R, Frueh BC. Effect of psychotherapy for depression via home telehealth on glycemic control in adults with type 2 diabetes: Subgroup analysis of a randomized clinical trial. J Telemed Telecare 2018 Oct;24(9):596-602. [doi: 10.1177/1357633X17730419] [Medline: 28945160]
42. Turvey C, Coleman M, Dennison O, Drude K, Goldenson M, Hirsch P, et al. ATA practice guidelines for video-based online mental health services. Telemed J E Health 2013 Sep;19(9):722-730. [doi: 10.1089/tmj.2013.9989] [Medline: 23909884]
43. Gonder-Frederick LA, Shepard JA, Grabman JH, Ritterband LM. Psychology, technology, and diabetes management. Am Psychol 2016 Oct;71(7):577-589. [doi: 10.1037/a0040383] [Medline: 27690486]
44. Kessler EA, Sherman AK, Becker ML. Decreasing patient cost and travel time through pediatric rheumatology telemedicine visits. Pediatr Rheumatol Online J 2016 Sep 20;14(1):54 [FREE Full text] [doi: 10.1186/s12969-016-0116-2] [Medline: 27646340]
45. Grossman JT, Frumkin MR, Rodebaugh TL, Lenze EJ. mHealth assessment and intervention of depression and anxiety in older adults. Harv Rev Psychiatry 2020;28(3):203-214 [FREE Full text] [doi: 10.1097/HRP.0000000000000255] [Medline: 32310834]
46. ICT use by households and individuals. Statistics Denmark. URL: https://www.dst.dk/da/Statistik/emner/uddannelse-og-viden/informationssamfundet/it-anvendelse-i-befolkningen [accessed 2022-03-02]
47. Home appliances 16. Statistics Denmark. URL: https://www.dst.dk/en/Statistik/emner/priser-og-forbrug/forbrug/elektronik-i-hjemmet [accessed 2022-03-03]

48. Koulil SS, Ferwerda M, van Beugen S, van Middendorp H, van de Kerkhof PCM, van Riel PLCM, et al. Tailored therapist-guided internet-based cognitive-behavioural treatment for psoriasis and rheumatoid arthritis: two case reports. Acta Derm Venereol 2018 Feb 07;98(2):225-233 [FREE Full text] [doi: 10.2340/00015555-2803] [Medline: 28952654]

Abbreviations

IT: information technology
PDSA: Plan-Do-Study-Act

Edited by A Kushniruk; submitted 31.05.21; peer-reviewed by H Mehdizadeh, A Joseph; comments to author 06.11.21; revised version received 01.01.22; accepted 03.01.22; published 21.03.22.

Please cite as:
Rothmann MJ, Mouritsen JD, Ladefoged NS, Jeppesen MN, Lillevang AS, Laustrup H, Ellingsen T
The Use of Telehealth for Psychological Counselling of Vulnerable Adult Patients With Rheumatic Diseases or Diabetes: Explorative Study Inspired by Participatory Design
JMIR Hum Factors 2022;9(1):e30829
URL: https://humanfactors.jmir.org/2022/1/e30829
doi:10.2196/30829
PMID:35311690

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org